# DIFFERENTIAL EXPRESSION OF CELLULAR GENES DURING A WEST NILE VIRUS INFECTION

**KOH WEE LEE**

**NATIONAL UNIVERSITY OF SINGAPORE**

**2004**

# DIFFERENTIAL EXPRESSION OF CELLULAR GENES DURING A WEST NILE VIRUS INFECTION

**KOH WEE LEE**
*(B.Sc.(Hons.), NUS)*

**A THESIS SUBMITTED**

**FOR THE DEGREE OF MASTER OF SCIENCE**

**DEPARTMENT OF MICROBIOLOGY**

**NATIONAL UNIVERSITY OF SINGAPORE**

**2004**

# MATERIALS FROM THIS STUDY HAVE BEEN PRESENTED AT THE FOLLOWING CONFERENCES

**WL Koh and ML Ng. 2003.** Differential regulatory profiles of West Nile virus-infected host cells. *6th Asia-Pacific Congress of Medical Virology*. Kuala Lumpur, Malaysia. (Excellence Award)

**WL Koh and ML Ng. 2004.** Global transcriptomic analysis of host cells with different susceptibility to West Nile virus infection. *11th International Congress on Infectious Diseases.* Cancun, Mexico. (ICID Scholarship)

**WL Koh and ML Ng. 2004.** Identification of potentially novel mechanisms involved in the pathogencity of West Nile virus. *5th Combined Scientific Meeting.* NUS, Singapore.

**WL Koh and ML Ng. 2004.** Insights into the mechanisms of cytopathic effects in host cells during a Flavivirus infection. *1st Pediatric Dengue Vaccine Initiative.* Bangkok, Thailand. (BMRC Travel Scholarship)

# ACKNOWLEDGEMENTS

I would like to express my sincere thanks and appreciation to the following people for their contributions during this study:

A/P Mary Ng – For her supervision and steadfast guidance during this trying period, and her support and time sacrificed in helping to produce this thesis, for which I owe my gratitude.

Loy Boon Pheng – For her efficient running of the lab, and her unwavering support in procuring materials. The induction rites were also memorable.

Li Jun, Bhuvana, Justin, John, Jason and all lab members – For their generous advice and support during periods of trials and tribulations.

Russell McInnes (Agilent Technologies) – For their prompt expert advice.

All family and friends – For their emotional support and encouragements during this wearisome period.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# SUMMARY

West Nile virus (WNV) is a mosquito-borne flavivirus and has the potential to cause fatal meningoencephalitis in infected victims. This re-emerging virus has recently caused large epidemics in the western hemisphere. Despite advances in WNV research, the mechanisms of cytopathology are still not known. Previous studies on WNV-host interactions have been limited. This area of research will be significant, as elucidations of these mechanisms will have direct implications in inhibiting the replication of the virus within the host. A screening of the global genomic expression was therefore carried out.

The initial studies on different human host cells have found that HeLa cells (cervical adenocarcinoma) were not as permissive as A172 cells (glial blastoma) to WN (Sarafend)V infection. An attempt to study the global transcriptomic profiles on host cells was subsequently carried out on two fronts: between virus-infected cells and mock-infected control cells, and between permissive cell lines and less-permissive cell lines. A time sequence study of the host response during the different phases of the virus infection was also carried out in A172 cells. Five time-points (1.5 h, 6 h, 12 h, 18 h, and 24 h) were carried out to cover the full spectrum of the virus replication cycle: from early to late phases of infection.

In the comparison between A172 and HeLa cells during a WNV infection, greater cytopathic effects accompanied with high virus titers were observed in the infected A172 cells. The intracellular levels of viral protein and RNA were quantified using immuofluorescence microscopy and quantitative PCR, respectively. Both virus components were consistently higher in A172 cells, and is therefore more permissive

to the virus infection. High-density microarray studies were utilized to elucidate the differences in host responses between the two types of cells. Four functional classes of genes belonging to cytoskeletal structure and functions, hexose metabolism, protein biosynthesis and RNA processing were found to be significantly differentially regulated between the two cell types. These classes of host responses could be responsible for the levels of permissiveness in the two cell lines.

In the time sequence study in WNV-infected A172 cells, differentially expressed genes during the course of the West Nile infection were clustered into 6 groups based on their gene expression patterns. Some of the functional groups that were differentially expressed at certain time points correlated with the stages in the virus replication cycle. These included the genes involved in the mitochondria, cellular transport and the endoplasmic reticulum. Further analysis is needed to understand the significance and impact of these genes on virus replication.

# INTRODUCTION

The Human Genome Project was launched about 10 years ago and the full sequence was recently published. This project has paved the way to the revolution in the life sciences that we are experiencing today. Its focus has started to shift gradually towards functional genomics, which deals with the functional analysis of genes and their products.

Techniques of functional genomics include methods for gene expression profiling at the transcript levels, protein levels, and bioinformatics. Among the techniques of functional genomics, both DNA microarrays and proteomics hold great promise for the study of complex biological systems with applications in molecular medicine (Celis *et al.*, 2000). These technologies are complementary, allowing high-throughput screening. In combination are expected to generate a vast amount of gene and protein expression data that may lead to a better understanding of the regulatory events involved in normal and disease processes. This could help to identify new networks of disease-associated alterations in humans.

Although much has been learned about the molecular biology of flaviviruses, there are still many unanswered questions. Since West Nile virus (WNV) alternates between insect vectors and vertebrates in nature, any cellular proteins that this virus uses during replication would be expected to be evolutionarily conserved. Of particular interest will be the identification of the cell protein(s) used for virus attachment and entry, and the elucidation of the molecular mechanisms involved in virus replication. Viruses use cell proteins during many stages of their replication cycles, including attachment, entry, translation, transcription/replication, and

assembly. Viruses also interact with cell proteins to alter the intracellular environment or cell architecture so that it is more favourable for virus replication. The replication can also inactivate intracellular defence mechanisms, such as apoptosis and interferon pathways. Mutations in the cell proteins involved can cause disruptions of these critical host-virus interactions. These virus-host interactions may thus represent novel targets for the development of new anti-viral agents.

Flavivirus-host interaction studies have not been extensive, and therefore, not well understood. Using West Nile (Sarafend) virus as a model for this study, an attempt was therefore made to elucidate the mechanisms of these virus-host interactions on a global scale.

# 1.0 LITERATURE REVIEW

## 1.1 Introduction to West Nile Virus

West Nile virus (WNV) is a mosquito-borne virus that was first isolated and identified as a distinct pathogen in 1937 from the blood of a febrile adult woman participating in a malaria study in the West Nile region of Uganda (Smithburn *et al.*, 1940). It was then classified as a flavivirus by a cross-neutralisation test (Calisher *et al.*, 1989; Wengler *et al.*, 1999). In the recent 76$^{th}$ Report of the International Committee on Taxonomy of Viruses (ICTV), they have assigned members of the genus into species (Heinz *et al.*, 2000; Mackenzie *et al.*, 2002). There are currently 27 mosquito-borne species, 12 tick-borne species and 14 species with no known vector. The appearance of the WNV in the United States in 1999 has increased interest not only in this virus, but also other flaviviruses, including dengue, yellow fever, Japanese encephalitis and tick-borne encephalitis viruses (TBEV).

## 1.2 West Nile Virus Epidemiology

The WNV isolates are grouped into two genetic lineages (1 and 2) on the basis of signature amino acid substitutions or deletions in their envelope protein (Berthet *et al.*, 1997). All WNV isolates that are associated with human diseases have been found in lineage 1, while lineage 2 viruses are mainly restricted to endemic enzootic infection in Africa (Jia *et al.*, 1999; Lanciotti *et al.*, 2002). Due to antigenic cross-reactivity between different flaviviruses, techniques such as *in situ* hybridization and sequence analyses of real-time polymerase chain reaction (PCR) products are required to unequivocally identify WNV as the causative agent in infections (Briese *et al.*, 2002; Lanciotti *et al.*, 2002).

In recent times, outbreaks have increased in frequency (Romania and Morocco in 1996; Tunisia in 1997; Italy in 1998; Russia and the United States in 1999; and Israel, France, and the United States in 2000) as well as the severity of the disease amongst those who developed clinical symptoms (Petersen and Roehrig, 2001). The WNV outbreaks in the USA have coincided with the emergence of a new variant of WNV designated "Isr98/NY99" that circulated in North America and the Middle East (Lanciotti *et al.*, 2002). This strain is characterized by a high avian death rate and an apparent increase in human disease severity (Solomon *et al.*, 2003).

This is consistent with the hypothesis that some changes in the neurovirulent properties of the virus had occurred (Ceccaldi *et al.*, 2004). In 2002, 39 states reported 4156 human WNV illness cases (O'Leary *et al.*, 2004), and the numbers increased to 9862 cases with 264 deaths in 2003 (CDC, 2004). The increased neurovirulence of Isr98/NY99 is accompanied by several novel modalities of transmission to humans, including transplacental transmission to the foetus, transmission via breast milk, blood transfusion, or laboratory contamination through percutaneous inoculation (Peterson and Roehrig, 2001; Hayes and O'Leary, 2004).

Wild bird species develop high levels of viremia after WNV infection and are able to sustain viremic levels of WNV of at least $10^5$ PFU/ml of serum (the minimum level estimated to be required to infect a feeding mosquito) for days to weeks. They are the main reservoir hosts in endemic regions for the virus, which can initiate epizootics outside the endemic areas (Bernard *et al.*, 2001; Petersen and Roehrig, 2001). The WNV has been isolated from *Culex, Aedes Anopheles, Minomyia, and Mansonia* mosquitoes in Africa, Asia, and the United States, but *Culex* species are the most

susceptible to infection with WNV (Burke and Monath, 2001; Ilkal *et al.*, 1997). Also *Culex* mosquitoes feed on wild bird species and they could have high levels of viremia (Turell *et al.*, 2000). Natural vertical transmission of WNV in *Culex* mosquitoes in Africa has been reported and is expected to enhance virus maintenance in nature (Miller *et al.*, 2000). Humans and horses are incidental hosts with low viremic levels and do not play a role in the transmission cycle.

Fever is the most common symptom observed in humans. The course of the fever is sometimes biphasic, and a rash on the chest, back, and upper extremities often develops during or just after the fever (Burke and Monath, 2001). Symptoms also include headaches, muscle weakness and disorientation. A portion of infected individuals develop encephalitis, meningoencephalitis, pan-meningo-polioencephalitis (Omalu *et al.*, 2003) or hepatitis. The brainstem, particularly the medulla, is the primary central nervous system (CNS) target. Humans aged 60 and older have an increased risk of developing fatal disease (George *et al.*, 1984; Sampson *et al.*, 2000; Chowers *et al.*, 2001). Flacid paralysis and muscle weakness, similar to polio-like syndrome, have also been reported in the absence of fever or meningo-encephalitis (Li *et al.*, 2003; Arturo *et al.*, 2003).

Histopathological studies after autopsy revealed that, although WNV could be detected in all major organs (spleen, liver, kidney, heart, etc.), most of the brains (88%) were also positive for viral antigens, including glial cells and neurones (Steele *et al.*, 2000). Neuropathogenicity was also observed in infected animals whereby poliomyeloencephalitis was characterized by T-lymphocytes and, to a lesser extent, macrophage infiltrates within the CNS, with multifocal glial nodules and some

nueronophagia (Cantile *et al.*, 2001). A Parkinson's disease-like syndrome, in which patients have mask-like faces, tremors and cogwheel rigidity is common in Japanese encephalitis (Misra and Kalita, 1997), correlating with the damage of the basal ganglia and thalamus. As high levels of WNV-reactive serum IgM antibodies could be detected in confirmed human cases of WNV encephalitis as late as 1.5 years after onset (Roehrig *et al.*, 2003) and also in animal studies (Xiao *et al.*, 2001), there is a possibility of viral persistence within the CNS.

## 1.3 Virus Morphology

West Nile virus belongs to the flavivirus family of viruses. The virions are small (~50nm in diameter), spherical, enveloped, and have a buoyant density of ~1.2g/cm$^3$. The WNV genome is a single-stranded RNA of positive polarity (mRNA sense) and is 11,029 bases in length, containing a single open reading frame (ORF) of 10,301 bases. The virus contain three structural proteins: a nucleocapsid protein (C protein, 14kDa), a lipid membrane protein (M protein, 8kDa), and a large envelope glycoprotein (E protein, 55kDa) carrying the majority of flavivirus antigenic and functional determinants (Heinz and Roehrig, 1990). The spherical nucleocapsid is ~25nm in diameter and is composed of multiple copies of the C protein. Cryo-electron microscopy data revealed that the virion envelope and capsid have icosahedral symmetry (Heinz *et al.*, 2000). The two viral envelope proteins, E and M, are both Type I integral membrane proteins with C-terminal membrane anchors (Mukhopadhyay *et al.*, 2003). Figure 1-1 shows the structure of the virus particle and Figure 1-2 shows the structural arrangement of the envelope proteins.

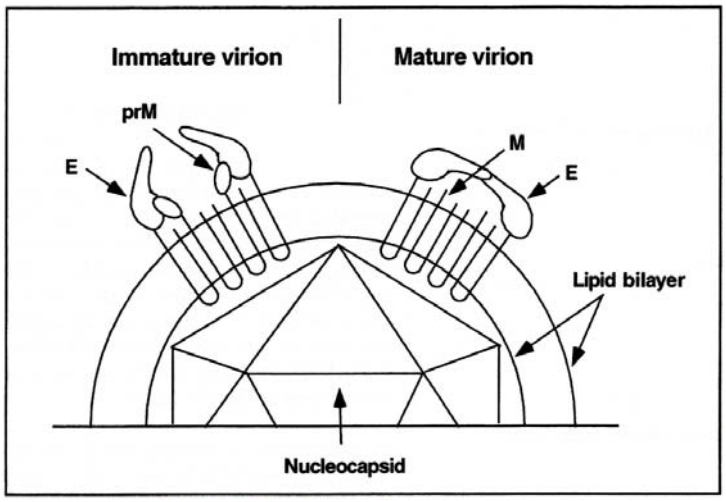



**Figure 1-1. The immature and mature flavivirus virions.** The heterodimers of prM and E are shown on the left (immature virion) and the homodimers of E, following cleavage of prM, on the right (mature virion). The icosahedral nucleocapsid consists of viral C protein and genomoic RNA, and is surrounded by a lipid bilayer in which the viral E and prM/M proteins are embedded. Viral maturation is triggered by the cleavage of prM to pr and M proteins by the host protease furin. (Shi, 2002)

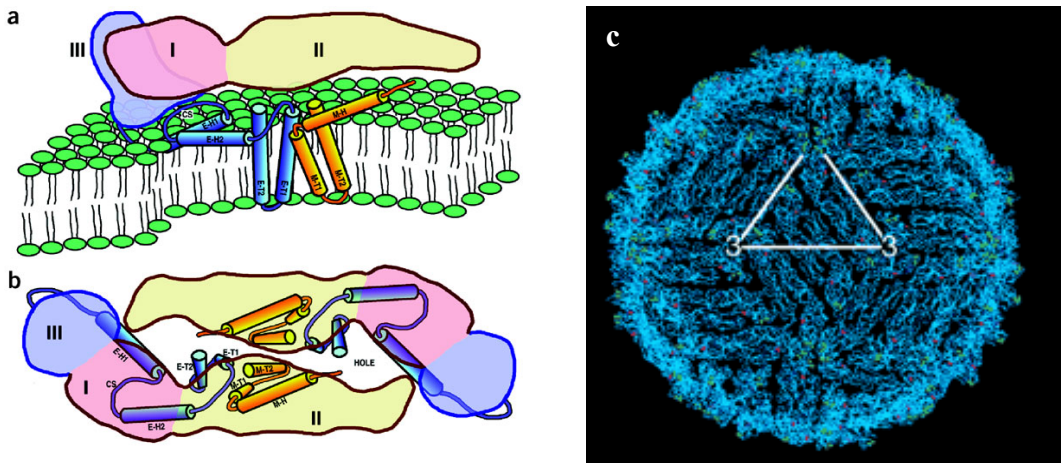



**Figure 1-2. Structural arrangement of flavivirus envelope protein.** (a) Diagrams of the flavivirus ectodomain and transmembrane domain proteins. The volume occupied by the ectodomain of an E monomer is pink (domain I), yellow (domain II) and lilac (domain III). (b) Homodimer arrangement of the E protein on the surface of the flavivirus particle (Zhang *et al.*, 2003). (c) Structure of the whole WNV with the homodimer E proteins arranged in a herringbone conformation (Mukhopadhyay *et al.*, 2003).

The recent determination of the structure of the entire virion of dengue virus type 2 by cryoelectron microscopy at a resolution of 24 Å has increased our understanding of the structure of the flavivirus virion. The structure has provided insights into the functions of its component parts (Kuhn *et al.*, 2002), especially with the elucidation of the crystal structure of surface glycoprotein E of TBEV by X-ray crystallography at 2 Å resolution (Rey *et al.*, 1995). The E glycoprotein is the principal stimulus for the development of neutralizing antibodies and contains a fusion peptide responsible

for inserting the virus into the host cell membrane. Generally, the E proteins of most flaviviruses are glycosylated, and the glycosylation of certain amino acid residues have been postulated to contribute to the pathogenecity of the virus (Beasley *et al.*, 2004). Varying N-glycosylation sites could also be important in epitope definition (Seligman and Bucher, 2003).

## 1.4 Virus Assembly and Maturation

WNV replicates in a wide variety of cell cultures, including primary chicken, duck and mouse embryo cells and continuous cell lines from monkeys, humans, pigs, rodents, amphibians, and insects, but does not cause obvious cytopathology in many cell lines (Brinton, 1986). It was demonstrated that although embryonic stem (ES) cells were relatively resistant to WNV infection before differentiation, they became permissive for WNV infection once differentiated, and die by the process of apoptosis (Shrestha *et al.*, 2003). Since flaviviruses are transmitted between insect and vertebrate hosts during their natural transmission cycle, it is likely that the cell receptor(s) they utilize is a highly conserved protein (Brinton, 2002). The receptor for WNV was found to be a 105-kDa protease-sensitive, N-linked glycoprotein in Vero and murine neuroblastoma 2A cells (Chu and Ng, 2003a), and was recently determined to be the αVβ3-integrin receptor (Chu and Ng, in press).

The pathway for flavivirus entry into host cells is through clathrin-mediated endocytosis, which is triggered by an internalization signal (di-leucine or YXXΦ) in the cytoplasmic tail of the receptor. Clathrin is assembled on the inside face of the plasma membrane to form an electron dense coat known as clathrin-coated pit. Clathrin interacts with a number of accessory protein molecules (Eps15, ampiphysin

and AP2 adapter protein) as well as the dynamin GTPase responsible for releasing the internalized vesicle from the plasma membrane (Marsh and McMahon, 1999).

This is followed by low-pH fusion of the viral membrane with the lysosomal vesicle membrane, releasing the nucleocapsid into the cytoplasm [(Heinz and Allison, 2000) (Fig. 1-3A)]. The reduced pH causes the conformational rearrangement of the E proteins, allowing the interactions of the virus E proteins with the lysosomal membrane to form hemifusion pores for the release of viral nucleocapsids into the cytoplasm for uncoating and replication (Modis *et al.*, 2004).

The RNA genome is released and translated into a single polyprotein (Fig. 1-3B). The viral serine protease, NS2B-NS3, and several cell proteases then cleave the polyprotein at multiple sites to generate the mature viral proteins (Fig. 1-3C). The viral RNA-dependent RNA polymerase (RdRp), NS5, in conjunction with other viral nonstructural proteins and possibly cell proteins, copies complementary minus strands from the genomic RNA template (Fig. 1-3D), and these minus-strand RNAs in turn serve as templates for the synthesis of new genomic RNAs (Fig. 1-3E). Upon WNV infection, extensive reorganization and proliferation of both smooth and rough endoplasmic reticula were observed (Ko *et al.*, 1979; Murphy, 1980; Westaway and Ng, 1980; Lindenbach and Rice, 1999). There were also induction of unique sets of membranous structures, but their functions during infection mostly remained elusive (Westaway *et al.*, 2002). One of such generic flavivirus-induced features, in both vertebrate and invertebrate cells, is the formation of vesicles packets that contains bi-layered membrane vesicles of 50-100 nm in size. These vesicles enclosed

distinctively single or double-stranded 'thread-like' structures during early stages of infection (Ng, 1987).

Flavivirus assembly occurs in association with the ER membranes (Fig. 1-2F, G). Intracellular immature virions, which contain heterodimers of E and prM, accumulate in vesicles and are then transported through the host secretory pathway [(Heinz *et al.*, 1994; Wengler, 1989) (Fig. 1-2H)]. It has been shown by electron microscopy that mature virions can be found within the lumen of endoplasmic reticulum (Matsumura *et al.*, 1977; Sriurairatna and Bhamarapravati, 1977; Hase *et al.*, 1989; Ng, 1987) at the perinuclear area of the cytoplasm (Murphy, 1980; Westaway and Ng, 1980). The glycosylated and hydrophilic N-terminal portion of prM is cleaved in the *trans*-Golgi network by cellular furin or a related protease (Stadler *et al.*, 1997). The C-terminal portion (M) remains inserted in the envelope of the mature virion (Murray *et al.*, 1993). The prM-E interaction may maintain the E protein in a stable, fusion-inactive conformation during the assembly and release of new virions (Heinz and Allison, 2000).

Assembly of WN (Sarafend) virus [WN(S)V] is, however, slightly different from the process shown above, which is generally true for other flaviviruses. With the use of cryo-immunoelectron microscopy, the precursor of nucleocapsid particles from WN(S)V was observed to be closely associated with the envelope proteins at the host cell's plasma membrane (Ng *et al.*, 2001). Instead of maturing within the endoplasmic reticulum, WN(S)V was found to mature (cis-mode) at the plasma membrane (Ng *et al.*, 1994). This contrasts with the trans-mode of maturation (Fig. 1-

2I) observed for most flavivirus where mature virus particles are released from cells by exocytosis (Mason, 1989; Nowak *et al.*, 1989).

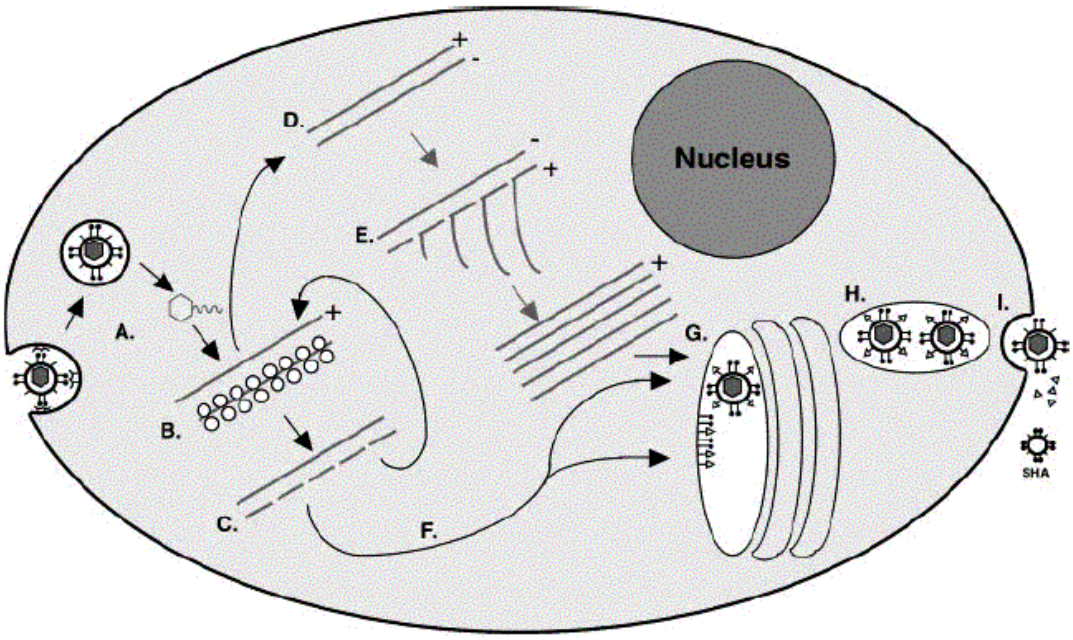



**Figure 1-3. The Flavivirus replication cycle.** *A*. Attachment and entry of the virion. *B*. Uncoating and translation of the virion RNA. *C*. Proteolytic processing of the polyprotein. *D*. Synthesis of the minus-strand RNA from the virion RNA. *E*. Synthesis of nascent genome RNA from the minus-strand RNA. *F*. Transport of structural proteins to cytoplasmic vesicle membranes. *G*. Encapsidation of nascent genome RNA and budding of nascent virions. *H*. Movement of nascent virions to the cell surface. *I*. Release of nascent virions. *SHA*, slowly sedimenting hemagglutinin, a subviral particle that is also sometimes released. (Brinton, 2002)

Egress of WNV had been observed to occur predominantly at the apical surface of polarized Vero cells, suggesting the involvement of a microtubule-dependent, polarized sorting mechanism for WNV proteins (Chu and Ng, 2002a). A recent study has shown that both E and C proteins were strongly associated and transported along the microtubules to the plasma membrane for assembly (Chu and Ng, 2002b). It was also observed in the same study that the association of E protein and microtubules was sensitive to high salt extraction but resistant to Triton X-100 and octyl glycoside extraction. This suggested that virus E protein and possibly also C protein associate effectively with the microtubules through an ionic interaction (Chu and Ng, 2002b).

## 1.5 Virus-Host Interactions

Infection and replication of viruses in vertebrate cells result in the alteration of expression of many cellular genes and these differentially expressed genes can be identified using a variety of techniques such as high-density DNA microarrays, differential display or subtraction hybridization (Manger and Relman, 2000). Such changes in host gene expression could be a cellular antivirus response, a virus-induced response that is beneficial or even essential for virus survival or a non-specific response that neither promotes nor prevents virus infection (Saha and Rangarajan, 2003).

Infection of diploid vertebrate cells with WNV has been reported to increase cell surface expression of MHC-1, which resulted from increased MHC-1 mRNA transcription activated by NF-κB (Kesson and King, 2001). Activation of NF-κB appeared to be mediated *via* virus-induced phosphorylation of inhibitor κB. Increased MHC-1 expression allows intracellular virus antigens to be presented, thus increasing the cell's susceptibility to virus-specific cytotoxic T-cell (CTL) lysis (Douglas *et al.*, 1994). This increase may also enhance tissue damage and immunopathology in an infected host (King *et al.*, 1993).

West Nile virus infection has also been reported to induce expression of non-conserved polymorphic intracellular adhesion molecule-1 [(ICAM-1)(CD54)] and its receptor, the integrin lymphocyte related antigen-1 [(LFA-1)(CD11a/CD18)] in infected cells (Shen *et al.*, 1995). The binding of ICAM-1 to its ligand has been found to increase the avidity of cellular conjugation between T cells and their target cells. This facilitates the interaction of antigen-targeted immune cells, and hence

contributing to more efficient antiviral responses. WNV-specific, interferon-independent induction of ICAM-1 was observed within 2 h after infection in quiescent but not replicating fibroblasts. The increase in MHC-1 and ICAM-1 expressions were found to be cell-cycle dependent, with up-regulation in $G_0$ phase compared to $G_1$ phase (Douglas *et al.*, 1994; Shen *et al.*, 1995). E-selectin (ELAM-1, CD62E), which is a rolling receptor for leukocyte adhesion, was found to increase maximally 2 h post-infection (p.i.), but declined to baseline levels within 24 h p.i. (Shen *et al.*, 1997).

Another common outcome of virus-host interaction is the physiological process of cell death. Apoptosis, which is an active and highly conserved process of cellular self-destruction with distinctive morphological and biochemical features, was observed in WNV-infected K562 and Neuro-2a cells and was shown to be *bax* dependent (Parquet *et al.*, 2001). Apoptosis has also been shown to be a major pathway of death in mouse neuronal cells infected with dengue virus (Despres *et al.*, 1996). Virus replication appears to be required since UV-inactivated virus failed to induce apoptosis. Apoptosis of cells might also be triggered by the M ectodomain (proapoptotic sequence) of WNV and similarly found in Dengue virus M protein (Catteau *et al.*, 2003). Since the introduction of WNV C protein into the nuclei of host cells has been shown to induce apoptosis, it could contribute to the pathogenesis of flavivirus infection (Yang *et al.*, 2002a). However, others found that neurons of mice infected with Murray Valley Encephalitis (MVE) virus do not show evidence of apoptosis, and the severity of the disease may be more linked to neutrophil infiltration and inducible nitric oxide synthetase activity in the CNS (Andrews *et al.*, 1999).

It was also reported that human umbilical vein endothelial cells (HUVEC) infected with WNV showed an increase in nitrite secretion and a rearrangement of zonula occludens-1 (ZO-1) and β-catenin (Wen *et al.*, 2001). It was thus postulated that WNV may modulate its entry into the CNS by altering cellular junctions of endothelial cells and leukocyte diapedesis across the endothelial cells.

The role of host genetic factors often has a part to play in the outcome of WNV infection. It was found that WNV replication was less efficient in cells that produce the normal copy of Oas1b as compared to those expressing the inactive mutated form (Lucas *et al.*, 2003). Variation in the response of individuals to flavivirus infection has also been observed in humans as well as in other host species. In mice, the alleles of a single Mendelian dominant gene, *Flv*, can determine whether an infection is lethal (Brinton, 1986) and segregates as a Mendelian dominant trait (Sangster *et al.*, 1993). The *Flv* resistance allele functions intracellularly to reduce the amount of virus produced, and the lower production of virus results in a slower spread of the virus in the host, both of which serve to give the host defence systems sufficient time to effectively eliminate the infection. Most of the currently used inbred mouse strains are susceptible. However BRVR, BSVR, C3H.RV, Det, PRI, and most wild *Mus musculus domesticus* are resistant.

As the severity of WNV infection varies between different individuals, it will be of interest to the study the role of host genetic factors and polymorphisms in WNV pathogenesis (Ceccaldi *et al.*, 2004). Other aspects of the host immune response may be critical in determining the outcome of human flavivirus infection. A role for innate immunity in JEV infection is suggested by the elevated IFNα levels found in plasma

and CSF (Burke and Morill, 1987). The humoral immune response to JEV, and to WNV infection is characterized by early production of IgM antibodies in both serum and CSF, followed by production of IgG (Martin *et al.*, 2002).

## 1.6 Neutralization of West Nile Virus Infection

The mouse model was used to point out the role of humoral immune response in limiting the spread of WNV infection in the CNS after primary replication in the lymph nodes (Diamond *et al.*, 2003a), and the role of CD8+ T cells in both recovery and immunopathology (Wang *et al.*, 2003). Recently, this model of infection has demonstrated that passive transfer of immune antibodies could improve the clinical outcome even after WNV had reached the CNS, although antibodies by themselves could not completely eliminate virus reservoirs host tissues (Engle and Diamond, 2003). Diamond and colleagues (2003b) have recently demonstrated the role of specific anti-WNV neutralizing IgM in preventing CNS infection and viral-induced death.

## 1.7 Global Genomic Analyses of Infected Host Cells

Within the past 5 years, increasing sophistication in infectious diseases research has caused an entirely new paradigm for fighting infectious disease to emerge. The unraveling of the genetic code of disease-causing microorganisms has allowed new methods to disrupt the disease process, which involve analysis of biological systems and molecular structures, thus producing a 'global picture' (Huang *et al.*, 2002). DNA microarrays lead the way in this area. This new technology allows researchers to

study how the entire genetic code of the invading microorganism interacts with the human cells it infects.

### 1.7.1 Microarrays

Microarrays consist of DNA molecules or probes, synthesized or attached to specific locations on a solid support, such as a coated glass surface. Arrays allow the identification of the sequence, and the abundance of each detected nucleic acid interrogated by the microarray. This is achieved by amplifying and labelling target nucleic acids from experimental samples and then monitoring the amount of label hybridized to each probe location (Schena, 2003).

The major types of DNA microarrays currently in use can be distinguished by the lengths of their probes and by the method of probe deposition onto hybridization substrates. Microarrays that carry sequences of more than ~100bp are commonly created using PCR products or cDNA clones, and are referred to as cDNA arrays. Microarrays that possess shorter DNA sequences are termed oligonucleotide microarrays (Southern, 2001).

Detection can be done by using radioisotopes like $^{32}P$, which gives precise quantification but has a wide shine and thus lower resolution. A common method is by using fluorescent labels like Cy5 and Cy3 that enables double labelling and high-resolution imaging (Southern, 2001), which is detected by using scanning confocal microscopy. In order to measure relative gene expression by using cDNA microarrays, RNA is prepared from the two samples to be compared, and labelled cDNA is made

by reverse transcription, incorporating either Cy3 (green) or Cy5 (red) fluorescent dye. The two labelled cDNAs are mixed and hybridized to the microarray, and the slide is scanned. In cases where the green Cy3 and red Cy5 signals are overlaid, yellow spots indicate equal intensity for the dyes. With the use of image analysis software, signal intensities are determined for each dye at each element of the array, and the logarithm of the ratio of Cy5 intensity to Cy3 intensity is calculated. Positive log (Cy5/Cy3) ratios indicate relative excess of the transcript in the Cy5-labelled sample, and negative log (Cy5/Cy3) ratios indicate relative excess of the transcript in the Cy3-labelled sample (Schena, 2003).

### 1.7.2 Microarray Applications

Gene expression microarray is a relatively new technology, yet it has already become a widely used tool in biology. The key fundamental issue of infectious diseases is how to globally and integratively understand the interactions between microbial pathogens and their hosts during infection (Huang *et al.*, 2002). Microarrays are ideally suited in this global approach.

#### 1.7.2.1 Expression Analyses: Gene Function and Elucidation of Regulatory Circuitry

Generally, gene expression experiments are designed to provide clues to gene product function, regulatory circuitry, and biochemical pathways. A gene is usually transcribed only when and where its function is required, determining the locations and conditions under which a gene is expressed. This allows inferences about its function towards pathogenesis. Experiments usually consist of comparing expression

levels in a disease tissue *versus* an unaffected tissue, or investigating cellular response in the presence and absence of an infectious agent (Warrington *et al.*, 2000).

The first application of global gene expression methods to pathogenesis used oligonucleotide arrays to monitor gene expression in primary human foreskin fibroblasts infected by human cytomegalovirus (CMV) and Toledo virus (Zhu *et al.*, 1998). The transcript abundance of 258 out of 6,600 human genes changed by more than fourfold, compared to uninfected cells, at either 8 or 24 h after infection. Some of these changes, such as induction of cytokines, stress inducible proteins, and many interferon-inducible genes, were consistent with induction of cellular immune responses (Zhu *et al.*, 1998).

With probe microarrays, the questions addressed are broader because thousands of genes are queried simultaneously, compared to the conventional methods of expression analyses of one or two genes per experiment. Large-scale analysis of the genome enables the elucidation of the expression patterns of the whole genome in a single experiment (Huang *et al.*, 2002).

#### 1.7.2.2 Expression Analyses: Pathogenesis

In addition to the simple observation of up and down regulation/expression of host genes, microarrays can also be used to ask very specific questions about the clinical manifestation of a disease and the role in pathogenesis of individual virulence factors (Huang *et al.*, 2002).

Transcription profiling of macrophages and epithelial cells infected by *Salmonella* confirmed increased expression of many proinflammatory cytokines and chemokines, signaling molecules, transcription activators and identified several genes previously unrecognized to be regulated by infection (Cummings and Relman, 2000). The macrophage study demonstrated that exposure to purified *Salmonella* lipo-polysaccharide (LPS) resulted in a very similar response profile to whole cells. The activation of macrophages with gamma interferon before infection modified the response. In epithelial cells, over-expression of κB (an inhibitor of NF-κB) blocked induction of gene expression for a number of regulated genes, underscoring the importance of NF-κB in the proinflammatory response (Detweiler *et al.*, 2001).

These data will help to identify genes with a critical role in pathogen progression and multiplication in the human host. Through the use of microarrays for monitoring gene expression profiles, infectomes of microbial and host cells during infection provide global and accurate information for building a comprehensive framework to interpret pathogenic processes (Huang *et al.*, 2002).

Global changes in gene expression of virus-infected cells in culture have been reported for several viruses such as human cytomegalovirus (Zhu *et al.*, 1998), herpes simplex virus (Mossman *et al.*, 2001), influenza virus (Geiss *et al.*, 2001a), Kaposi's sarcoma associated virus (Renne *et al.*, 2001), human papillomavirus (Chang and Laimins, 2000) and human immunodeficiency virus type 1 (Geiss *et al.*, 2001b). Studies on neurotropic viruses include rabies virus (Prosniak *et al.*, 2001) and Sindbis virus (Johnston *et al.*, 2001).

A virus-host interaction study on dengue virus, which is another *flavivirus*, was recently carried out using Affymetrix microarrays on human umbilical vein endothelial cells (Warke *et al.*, 2003). They found 269 genes that were induced and 126 genes that were repressed. Broad functional responses that were activated included the stress, defense, immune, cell adhesion, wounding, inflammatory, and antiviral pathways. In another study, a microarray study was conducted on the pathogenic WNV (NY strain) which was observed to evade the host cell innate antiviral response (Fredericksen *et al.*, 2004). However, this was carried out on 293 cells (human epithelial kidney) which do not represent the natural CNS hosts. Nevertheless, Fredericksen and colleagues (2004) reported that the WNV was able to replicate efficiently despite the activation of IFN-β and several IFN-stimulated genes late in infection through the IFN regulatory factor 3 (IRF-3) pathway.

#### 1.7.2.3 Expression Analyses: Time-Course Study

A similar experimental design has been used to examine the global effects of HIV-1 infection on cultured CD4-positive T cells. One study concluded that HIV-1 infection resulted in differential expression of 20 of the 1,506 human genes monitored and that most of these changes occurred only after 3 days in culture (Corbeil *et al.*, 1999). In contrast, the preliminary results of an independent study using a similar design indicated that substantial HIV-induced transcription changes began very early after inoculation (Geiss *et al.*, 2000). The latter study confirmed activation of nuclear factor-κB (NF-κB), p68 kinase, and RNase L.

A time-course study of *Cryptococcus neoformans* infection of human brain microvascular endothelial cells (HBMEC) was done using oligonucleotide microarrays to monitor the infectomes of 12,558 human genes. An ontology (gene functional classification) analysis revealed gene expression patterns of different subsets of genes within the same functional class. For example, among the 7 time-point samples, the changes in expression profiles of the 29 MHC class II-related genes suggested that *C. neoformans* may contain superantigens stimulating the immune system (Huang *et al.*, 2002).

### 1.7.3 Microarray Data Management and Manipulations

Microarray experiments churn out massive amounts of data in a single experiment and analyzing the data has proven to be more complex than carrying out the experiment itself. This is made especially more daunting as a standardized approach to analyzing microarray data is not present (Nadon and Shoemaker, 2002). Microarray data are cumbersome with hands-on data transformation, leading to human errors which often have dramatic consequences and thus, altering results (Grant *et al.*, 2003). Data loading and storage usually involves several parsing and data transportation steps, each of which can corrupt the data from their original state. Data integrity management is therefore important in preventing data corruption.

#### 1.7.3.1 Identification of Differentially Regulated Genes

To identify genes that are up- or down-regulated in the sample compared to control, scaling of the data is first required (Knudsen, 2002). Normalization is carried out to ensure that the expression levels in the sample are comparable to the expression

levels in the control. There are a number of reasons why data must be normalized. This includes the unequal quantities of starting RNA, differences in labelling or detection efficiencies between the fluorescent dyes used, and systematic biases in the measured expression levels (Quackenbush, 2002). The $\log_2$ (ratio) values can have a systematic dependence on intensity, which most commonly appears as a deviation from zero for low-intensity spots. Locally weighted linear regression (lowess) analysis has been proposed as a normalization method that can remove such systematic biases or intensity-dependent effects in the $\log_2$ (ratio) values. Lowess uses a weight function that deemphasies the contributions of data from array elements that are far from each point (Yang *et al.*, 2002b).

Duplication is essential for identifying and reducing the variation in any experimental assay. Duplication in a two-colour spotted array experiment can be carried out by a dye-reversal or flip-dye analysis for each RNA sample (Churchill, 2002). This process may help to compensate for any biases that may occur during labelling or hybridization; for example, if some genes preferentially label with the red or green dye. Experimental variation during duplication will lead to a distribution of the measured values for the log of the product ratios, $\log_2(T1i*T2i)$. The consistent array elements between a flip-dye duplicates are expected to have a value for $\log_2(T1i*T2i)$ close to zero. Inconsistent measurements have a value 'far' from zero and can be eliminated from further consideration. The stringency of this elimination can be chosen based on the number of standard deviation of the mean. Averaging over the duplicates will then reduce the complexity of the data set (Quackenbush, 2002).

Differentially regulated genes or genes exhibiting the most significant variation are often identified using a fixed fold-change cut-off (generally twofold) from the $\log_2$(ratio) figures. Another more sophisticated approach involves calculating the mean and standard deviation of the distribution of values and defining a global fold-change difference and confidence; this is essentially equivalent to using a *Z*-score for the data set. Using a sliding window to determine the local structure of the data set, one can calculate the mean and standard deviation within a window surrounding each data point. An intensity-dependent *Z*-score threshold is defined to identify differential expression, where *Z* simply measures the number of standard deviations a particular data point is from the mean (Yang *et al.*, 2002c). Differentially expressed genes at the 95% confidence level would be those with a value of more than 1.96 standard deviations from the local mean. At higher intensities, this allows smaller changes to be identified, while applying more stringent criteria at intensities where the data are naturally more variable at the lower intensity regions.

#### 1.7.3.2 Identification of Gene Expression Patterns

The data from expression arrays is often of a high dimensionality. A 10 array experiment with 15,000 genes will constitute a matrix of 10 x 15,000. To facilitate a visual analysis of the data, a reduction of the dimensionality of the matrix is necessary (Knudsen, 2002). Since visual analysis is traditionally performed in two dimensions, clustering algorithms can help in this process by grouping significantly changed genes into clusters that behave similarly under different conditions.

The object of hierarchical clustering algorithm is to compute a dendrogram that assembles all elements into a single tree. For any set of *n* genes, an upper-diagonal

similarity matrix is computed, which contains similarity scores for all pairs of genes. The matrix is scanned to identify the highest value, representing the most similar pair of genes. A node is created joining these two genes, and a gene expression profile is computed for the node by averaging observation for the joined elements. The similarity matrix is updated with this new node replacing the two joined elements, and the process is repeated *n*-1 times until only a single element remains (Eisen *et al.*, 1998). A graphical representation of the primary data is obtained by representing each data point with a colour that quantitatively and qualitatively reflects the original experimental observations. The end product is a representation of complex gene expression data that, through statistical organization and graphical display, allows biologists to assimilate and explore the data in a natural intuitive manner. Relationships among objects (genes) are represented by a tree whose branch lengths reflect the degree of similarity between the objects. Such methods are useful in their ability to represent varying degrees of similarity and more distant relationships among groups of closely related genes (Eisen *et al.*, 1998).

Hierarchical clustering fails when the number of genes reaches several thousands. Calculating the distances between all of them becomes time consuming. Removing genes that show no significant change in any experiment is one way to reduce the problem. Another way is to use a faster algorithm, like *K*-means clustering (Knudsen, 2002).

In *K*-means clustering, the number of clusters can be decided by the user, and the algorithm then randomly assigns each gene to one of the *K* clusters. The distance between each gene and the centre of each cluster (centroid) is calculated, and the

genes are continually shifted to the closest cluster. The centroids will be recalculated after each step and the algorithm will stop after the cluster centroids no longer change (Soukas *et al.*, 2000). The Figures of Merit (FOM) algorithm can be used to determine the appropriate number of clusters for *K*-means clustering. A FOM is an estimate of the predictive power of a clustering algorithm. It is computed by removing each experiment in turn from the data set, clustering genes based on the remaining data, and calculating the fit of the withheld experiment to the clustering pattern obtained from the other experiments. The lower the adjusted FOM value is, the higher the predictive power of the algorithm (Yeung *et al.*, 2001).

Another method of clustering is the Self Organizing Tree Algorithm (SOTA). This involves the use of unsupervised neural network, which grows by adopting the topology of a binary tree. The result of the algorithm is a hierarchical cluster obtained with the accuracy and robustness of a neural network. Since SOTA runtimes are approximately linear with the number of items to be classified, it is especially suitable for dealing with huge amounts of data (Herrero *et al.*, 2001).

The *t*-test is used to determine if genes are significantly different from a pre-defined mean value. Each gene whose mean $\log_2$ expression ratio over all experiments is significantly different from the mean value of zero (i.e. no change in expression) is assigned to one cluster. *T*-values are calculated for each gene, and *p*-values are computed either from the theoretical *t*-distribution, or from permutations of the data for each gene. The user determines the critical p-value to determine significance (Pan, 2002).

#### 1.7.3.3 Quantitative Real-Time PCR (qRT-PCR) to Quantify Transcript Levels

The two commonly used methods to analyze data from qRT-PCR experiments are absolute quantification and relative quantification. Absolute quantification determines the input copy number, usually by relating the PCR signal to a standard curve. This is performed when it is necessary to determine the absolute transcript copy number. Relative quantification relates the PCR signal of the target transcript in a treatment group to that of another sample such as an untreated control, and is sufficient when only the relative change in gene expression is needed. The $2^{-\Delta\Delta C}{}_T$ method of analysis is often used to calculate the relative changes in gene expression (Livak and Schmittgen, 2001).

Standard curves derived from serial dilutions of samples provide a useful tool to evaluate the consistency of the PCR reactions. This will help to test the response of the reagent system to different starting quantities that may be found in the test samples. The assay should return predictable and consistent results based on the inputs, similar to a mathematical formula. $R^2$ is the correlation coefficient squared and is a measure of how closely the calculated CT values fit the expected values. $R^2$ is a positive number, and the closer to 1.00, the better the fit (BioRad, 2004). If the points on a standard curve do not fall on a straight line, it might be the result of some kind of inhibitor present in the test sample, and is representative of standard curves with $R^2$ values above 1.00. The inhibitor is diluted out at lower concentrations, so it does not affect the kinetics of the experiment at these concentrations, which may be introduced during the various steps of the cDNA isolation process (BioRad, 2004).

## 1.8 Objectives

There are three general aims of this study:

a. To optimize the techniques for genomic microarray studies that are tailored for virus-host interactions, as well as subsequent downstream confirmatory tests.
b. To identify groups of cellular genes that might be important for the pathogenesis of WNV infection by comparative analysis of permissive and less permissive cells.
c. To carry out a time-course study from early- to late-phase infection to determine the changes in gene regulation in response to virus replication.

# 2.0 MATERIALS AND METHODS

## 2.1 Cell Culture

### 2.1.1 Tissue Culture Techniques

All solutions and media for cell culture were made using type 1 grade reagent water (NANOpure, Barnstead, USA). The chemicals used were also of ultrapure grade. Glass bottles were used for storage of the media. These have screw-capped lids with non-toxic plastic blue washer. All cell culture and media preparations were done under aseptic conditions in a laminar flow hood (Gelman Sciences, Australia) or in a biohazard hood (Gelman Sciences, Australia). Cells used in this study were grown in sterile 75cm$^2$ plastic tissue culture flasks from Nunc (Denmark).

### 2.1.2 Cell Lines

Two human cell lines were used in this study. They were the HeLa cell line and the neuroblastoma cell line, A172. HeLa cells were derived from cervical adenocarcinoma cells obtained from a 31 years old Negroid woman. A172 cells were derived from the glioblastoma brain tumour cells of a 53 year old male (Giard *et al.*, 1973). The passage number of the cell lines used in this study was between 50 and 80. Vero cells, a continuous cell line that was derived from African green monkey kidney cells, were used for propagation of the WN virus. The passage number of the Vero cell line used was between 50 and 80.

### 2.1.3 Media for Cell Culture

Dulbecco's Modified Eagle's media [(DMEM) (Sigma, USA – Appendix 1a)] was used as the growth media to culture both HeLa and A172 cells and Medium 199 [(M199)(Sigma, USA – Appendix 1b)] was used to culture Vero cells. The media was prepared to manufacturer's specifications, which was supplemented with 10% foetal calf serum (FCS). Sodium bicarbonate was added as a buffering agent, and the pH of the media was adjusted to 7.2.

### 2.1.4 Regeneration, Cultivation and Propagation of Cell Lines

Cells in cryo-vials were stored in liquid nitrogen. To revive the cells, each vial of the desired cell line was retrieved from liquid nitrogen storage and immediately thawed in a 37°C water bath. When thawed, the cells were transferred into a 75 $cm^2$ culture flask and 15 ml of DMEM was added. The growth media was needed to dilute the toxic effects of dimethysulphoxide (DMSO), which was present in the preserving medium. The cells in the flasks were then incubated at 37°C. The growth media was decanted after 12 h and replaced with fresh media, after which the cells were allowed to grow to confluency for about 3-4 days.

When the cells were confluent, the medium was discarded and the cell monolayer was rinsed once with 10ml PBS (Appendix 1c). This was followed by the addition of 3ml trypsin-versene solution (Appendix 1d) and incubated at 37°C for about 2 min. It was then observed under microscope to ensure that the cells have rounded up. The flask was tapped gently to dislodge the cell monolayer. Five ml of growth medium was immediately added to inactivate the enzymatic effect of the trypsin-versene solution.

The cell aggregates were resuspended by pipetting up and down for a few times. The suspension of HeLa or A172 cells was split into a ratio of 1:4 for seeding into 75 $cm^2$ culture flasks, and topped up to 17 ml with growth media. The cells were cultivated at 37°C, in a humidified 5% $CO_2$ incubator (Lunaire, USA). The monolayer reached confluency in about 3 days and was used for subsequent experiments.

### 2.1.5 Cultivation of Cells on Coverslips

Cells were grown on coverslips for immunofluorescence microscopy. Glass coverslips of diameter 13 mm (ARH, UK) were washed with 90% ethanol for 30 min and then boiled in double-distilled water for about 10 min. The coverslips were then left to air dry. Dry sterilisation was done in a hot air oven at 160°C (Jouan, USA) for 2 h. The individual coverslips were subsequently placed aseptically in a 24-well tissue culture tray (Nunc, Denmark). Cells from a confluent cell monolayer in a 75 $cm^2$ flask were used. The monlayer was trypsinized (Section 2.1.4) and resuspended in 30 ml of DMEM. Two ml of the cell suspension were dispensed into each well. The trays were then left at 37°C in the 5% $CO_2$ incubator (Lunaire, USA) until they were 50 to 70% confluent.

## 2.2 Infection of Cells

### 2.2.1 Virus Strains

The virus used in this study was West Nile (Sarafend) virus [WN(S)V] – a gift from Emeritus Professor Westaway, Sir Albert Sakzewski Virus Research Laboratory, Queensland, Australia. The virus stock was propagated in Vero cells throughout the study, and introduced into the human cell lines, HeLa and A172, for infection studies.

The virus was not 'adapted' to the human cell lines prior to infection, so as to ensure that a base level of comparison can be obtained by using the same virus stock. This was also to prevent any form of attenuation to the virus when grown in HeLa cells (Dunster *et al.*, 1990).

### 2.2.2 Infection of Cell Monolayers and Production of Virus Pool

A confluent cell monolayer of about 3 days old from a 75 $cm^2$ culture flask was used for infection. The growth medium was discarded and the monolayer was washed with 5 ml of virus diluent (Appendix 2a). A volume of 1 ml of virus suspension with multiplicity of infection (MOI) = 10 was inoculated onto the cell monolayer. The flask was incubated at 37°C for 1 h and rocked every 15 min to ensure even infection of the cell monolayer. After 1 h of virus adsorption, 10 ml of maintenance medium (Appendix 2b, c) was added to the flask. The infected cells were then incubated at 37°C for 24 h. Mock-infected controls on HeLa and A172 cells were prepared in the above manner; with 1 ml of virus diluent (Appendix 2a) used instead of virus suspension.

At the end of the incubation period, the maintenance medium containing extracellular virus particles was then harvested. The supernatant was first spun on a bench top centrifuge (Sigma Model 3K15, USA) at 1,000 rpm for 10 min at 4°C to remove cell debris. One ml of this supernatant was aliquoted into sterile cryo-vials, sealed and frozen immediately in cold ethanol (-80°C). The vials were subsequently stored at -80°C.

### 2.2.3 Plaque Assay

Virus stock was diluted in ten-fold serial dilutions using virus diluent (Appendix 2a) from $10^{-1}$ to $10^{-8}$ dilutions. Aliquots containing 0.1 ml of the appropriate dilutions were inoculated onto a day-old confluent Vero cell monolayer grown in a 24-well culture plate (Nunc, Denmark). The virus was allowed to adsorb to the cells at 37°C for 1 h, with gentle rocking at 15 min intervals. Following that, the excess inoculate were removed and the wells washed gently with virus diluent. One ml of overlay media (Appendix 2d) was added to each well. The plate was placed in a humidified 37°C, 5% $CO_2$ incubator (Lunaire, USA). After about 4 days of incubation, the overlay media was decanted and then stained with 1% crystal violet solution (Appendix 2e) for 1 h at room temperature. Thereafter, the plate was rinsed twice with water and dried. The number of plaques obtained was then counted.

The virus was plaqued on Vero cells, even though they had been passaged in HeLa and A172 cells, so that a basal level of comparison can be obtained. It had also been reported that HeLa cell plaque assays were unreliable (Dunster *et al.*, 1990).

## 2.3 Light Microscopy

When the monolayers reached 70% confluency, the cells were infected with WN(S) virus as before (Section 2.2.2). The flasks were incubated for 24 h until cytopathic effects (CPE) was observed. The flasks were then visualised under an optical microscope (IX81, Olympus, Japan) that was linked to a digital camera. Pictures of the virus-infected and mock-infected control cells were taken under phase-contrast at magnifications of 100x and 400x.

## 2.4 Genomic Expression Studies

Microarray assays are essentially miniaturized hybridization assays for studying thousands of nucleic acid fragments simultaneously. Figure 2-1 shows the key components of a basic microarray experiment.

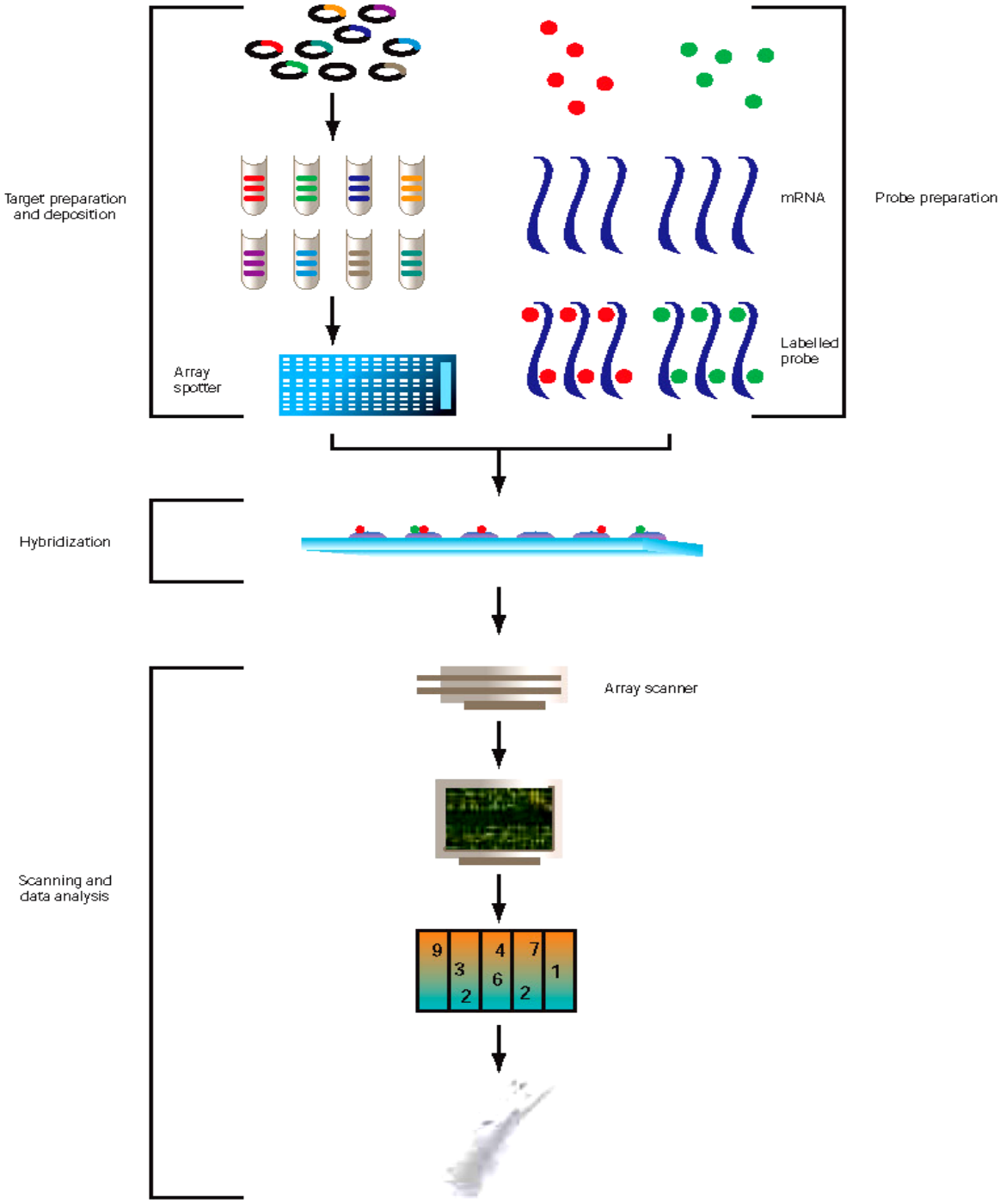



**Figure 2-1.** The main steps in a microarray experiment involves probe preparation, hybridization, scanning and data analysis. (Amersham Biosciences, 2002)

### 2.4.1 Microarrays

Microarrays consist of a collection of nucleic acid sequences immobilized onto a solid support so that each unique sequence forms a tiny feature, called a 'spot' or 'target'. Agilent's Human 1A Oligo Microarrays (Agilent Technologies, USA) were used for this study. It comprised of 22,575 (60-mer) oligonucleotide probes representing 17,803 well-characterized, full length, human genes from the Incyte Genomics Foundation Database. One thousand and seventy-five of the probes were quality control features consisting of negative controls, performance controls, oligo synthesis controls, microarray location controls and feature morphology controls.

### 2.4.2 Probe Labelling

Total cellular RNA was the starting material for this microarray experiment, which was subsequently converted to a labelled population of cDNA, known as the 'probe'. These probes frequently consisted of several thousands of different labelled fluorescent cyanine dyes (Cy3 and Cy5) coupled to the nucleic acid fragments.

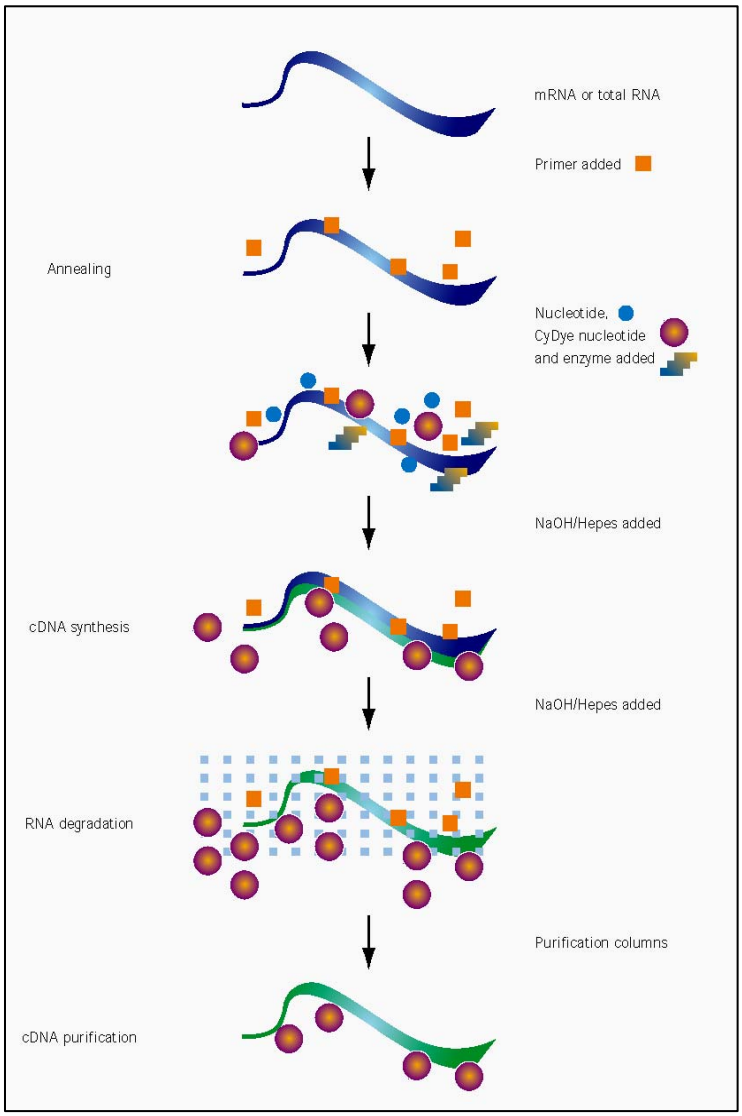



**Figure 2-2.** The main steps involved in probe labelling.

These fluorescent dyes are compatible with current microarray formats, with high spectral separation, high incorporation rates with a variety of enzymes, and fluoresce brightly when dry. Fluorescence has the advantage of permitting the detection of two or more different signals in one experiment. It has also increased the accuracy and throughput of microarray analysis over filter-based macroarrays, in which only one radioactively labelled sample can be conveniently analyzed at a time (Amersham Biosciences, 2002). CyDyes should be protected from light during all handling and storage. Figure 2-2 shows the main steps involved in probe labelling.

However, it should be noted that CyDyes exhibit different quantum yields during incorporation. Cy5 also has the disadvantage that it sometimes gives high background levels on glass surfaces and is more sensitive to photobleaching than Cy3 [(Bilban *et al.*, 2002) (See also Section 4.2.1.2)].

#### 2.4.2.1 Total RNA Isolation from Cell Culture

RNA is prone to disintegration from ubiquitous ribonucleases (RNase), therefore it is important to stabilize RNA and adopt proper RNA handling techniques. A cell monolayer was first scrapped off using a cell scrapper (Nunc, Denmark) and pelleted in RNase-free eppendorf tubes by centrifugation at 500 x *g* for 5min. The supernatant was then discarded and the cells were washed with PBS (Appendix 1c) to remove all media. The cells were resuspended in 100 µl of PBS and 1 ml of RNA*later* RNA Stabilization Reagent (Ambion Inc, USA) was added, which protects RNA in cells by preventing unwanted changes in the gene-expression patterns due to RNA degradation or new induction of genes. The cell sample can then be stored at 4°C.

Total RNA isolation was carried out using QIAGEN RNeasy Mini Kit (QIAGEN GmbH, Germany) according to the manufacturer's recommended protocol. Briefly, the cells were pelleted at 3000 x *g* to remove the RNA*later* RNA Stabilization Reagent. A higher centrifugal force was necessary to pellet the cells since the stabilization reagent has a higher density that most cell-culture media. After the lysis of cells to release RNA, homogenization of the sample is required to reduce the viscosity of the cell lysates by shearing the high-molecular weight genomic DNA and other high-molecular weight cellular components to create a homogeneous lysate. Homogenization would disrupt the cells and thus increase the yield of RNA and this was carried out using QIAshredder (QIAGEN GmbH, Germany) according to the manufacturer's protocol. RNA was eluted out in 60 µl of RNase-free water and stored at -20°C for later use.

#### 2.4.2.2 RNA Quantification and Quality Determination

RNA concentration, of an appropriately diluted sample in DEPC water (Appendix 3a), can be determined by measuring the absorbance at 260 nm ($A_{260}$) in a spectrophotometer (Shimadzu-UV 1601, Australia). An absorbance of 1 unit at 260 nm corresponds to 40 µg of RNA per ml. Cuvettes (Hellma GmbH, Germany) used for measurement had to be RNase-free. The purity of RNA was determined by taking the ratio of the readings at 260 nm and 280 nm ($A_{260}/A_{280}$). Pure RNA has an $A_{260}/A_{280}$ ratio of 1.9-2.1.

#### 2.4.2.3 Determination of RNA Integrity

The integrity and size distribution of total RNA extracted was checked by denaturing formaldehyde-agarose (FA) gel (Appendix 3b) electrophoresis. The respective ribosomal bands (1.9kb and 5.0kb for 18S and 28S rRNA, respectively) should appear as sharp bands on stained gels. Degraded RNA samples (smearing of the ribosomal bands) should not be used to proceed with microarray analysis. Prior to running the gel, equilibrate the gel in 1 X FA gel running buffer (Appendix 3c) for at least 30 min. Two µl of RNA sample was mixed with 8 µl of RNA loading buffer (Appendix 3d) and mixed. Ten µl of each mixture was incubated for 5 min at 65°C, and then chilled on ice. The equilibrated FA gel was electrophoresied at 100V for 1.5 h. The gel was subsequently visualized under UV and images captured using ChemiGenius$^2$ (Syngene, UK).

#### 2.4.2.4 Reverse Transcription and Labelling

mRNA has to be reverse-transcribed into labelled cDNA for use on microarrays. In the initial phase of microarray experiments, CyScribe First-Strand cDNA Labelling Kit (Amersham Biosciences, USA) was used to incorporate Cy3-dCTP and Cy5-dCTP (both from Amersham Biosciences, USA) into cDNA probes in first-strand cDNA synthesis reactions. This labelling kit was used in microarray experiments for the comparison between A172 and HeLa cells after 24 h p.i. with WN(S)V. Twenty-five µg of total RNA from each infected and mock-infected cell lines was used as the starting sample. Priming with anchored oligo(dT) will direct the start of the synthesis of cDNA from the 5' end of the poly-A-tail by the reverse transcriptase. The labelled cDNA obtained using this method undergoes no amplification in amount.

The fluorescent cDNA probes require purification from the RNA template, and unincorporated fluorescent nucleotides have to be removed, in order to maximize hybridization signal and minimize non-specific hybridization background on microarrays. The RNA template was degraded with alkaline hydrolysis treatment with the addition of 2 µl of 2.5 M NaOH (Appendix 3e) into each microcentrifuge tube containing the labelling reactions and incubated at 37°C. Ten µl of 2 M HEPES free acid (Appendix 3f) was added to each reaction tube after 15 min to neutralize the reaction mixture.

After neutralization, the labelling reactions were ready for purification. This was carried out using CyScribe GFX Purification Kit (Amersham Biosciences, USA) according to the manufacturer's protocol, which achieved 50% or higher of cDNA recovery and removed unincorporated nucleotides from the labelled cDNA. The purified labelled cDNA was eluted out in 120 µl of Elution Buffer.

#### 2.4.2.5 Quantification of cDNA Yield and Incorporation of Fluorescent Nucleotides

The yield of fluorescent probe is determined not only by the success of the labelling step and by the amount of template RNA used in the reaction, but also by the recovery of labelled material from the purification system. In order to produce reproducible and high-quality microarray results, it is imperative to use balanced and optimal amounts of fluorescently-labelled samples in hybridizations. Quantifying the amount of Cy3 and Cy5 incorporation into cDNA is therefore necessary.

Absorbance of the purified sample was measured against blank ($ddH_20$) at 550 nm for Cy3 and at 650 nm for Cy5 in a 100 µl cuvette (Hellma GmbH, Germany). The amount of Cy3 and Cy5 incorporated into cDNA was calculated as follows:

$$p\text{mol CyDye in sample} = (A/E) \times Z \times 10^6$$

where A = absorbance of Cy3 (at 550 nm) or Cy5 (at 650 nm), E = extinction coefficient for Cy3 (150,000 $lmol^{-1}cm^{-1}$) or Cy5 (250,000 $lmol^{-1}cm^{-1}$), *Z* = volume of purified cDNA in µl (=120 µl).

### 2.4.3 Microarray Hybridization

In microarray hybridization, the labelled fragments in the probe are expected to form duplexes with their immobilized complementary targets. This requires the nucleic acids to be single-stranded and accessible to each other. The number of duplexes formed during hybridization reflects the relative number of each specific fragment in the probe. Hybridization should be carried out under stringent conditions (high temperature) that do not promote annealing of non-complementary fragments.

Hybridization preparations were carried out according to Agilent Technologies (USA) Oligonucleotide Microarray Hybridization protocol in conjunction with Agilent's In-situ Hybridization Kit Plus (Agilent Technologies, USA). Equal amounts of Cy3 and Cy5 labelled cDNA (20-40 *p*mol each) were suspended in 100 µl of nuclease-free water (Gibco BRL, USA). The resuspended cDNA was heat denatured for 3 min at 98°C, and then cooled to room temperature. Twenty-five µl of 10x Control Targets and 125 µl of 2x Hybridization Buffer were added to give a total volume of 250 µl, and immediately used for hybridization.

Lucidea SlidePro Hybridizer (Amersham BioSciences, USA), an automated instrument for performing hybridization, was initially used in the hybridization reactions for the comparison between A172 and HeLa cells. An automated system helps to minimize and control variations in environmental conditions during hybridization to produce a consistent fluorescent signal within and between slides. Variations are a common occurrence in manual hybridization methods due to probe depletion and target saturation. Pre-wash of the instrument was carried according to the manufacturer's protocol before inserting the slide. Two hundred and fifty µl of the reaction mixture was injected into the chamber and allowed to hybridize to the target for 17 h at 60°C. After hybridization, the slides are washed in Wash Solution 1 (Appendix 3g) for 10 min at room temperature to remove all unattached and loosely bound probe molecules, and subsequently in Wash Solution 2 (Appendix 3h) for 5 min at 4°C. The slides were then dried immediately with an air-pump to prevent smearing. The hybridization, washing and drying steps were automatically performed by the Lucidea SlidePro.

### 2.4.4 Scanning

Scanning of the processed microarrays was performed on Axon's GenePix 4000B Microarray Scanner (Axon Instruments, USA). It consists of a confocal microscope attached to a detector system and scanned with two different lasers that emit light at 635 nm and 532 nm for Cy5 and Cy3, respectively, allowing high-resolution (5 µm) detection of the hybridization signals. Scanning was carried out according to the manufacturer's user guide before undergoing data analysis.

### 2.4.5 Protocol from Agilent Technologies (USA)

After the scanning of the initial lot of microarray slides that were labelled with Amersham's CyScribe First-Strand cDNA Labelling Kit and hybridized on the Lucidea SlidePro Hybridizer, it was observed that the hybridization pattern on some of the slides were 'patchy'. The amount of fluorescence from the spots was not evenly distributed across the whole microarray slide. This resulted in wastage of some microarray slides as they could not be reused. An alternative probe labelling and hybridization protocol was therefore utilized to optimize the results. These were carried out on the second phase of microarray experiments, which was a time-series of virus infections (i.e. 1.5 h, 6 h, 12 h, 18 h, 24 h p.i.) in A172 cells.

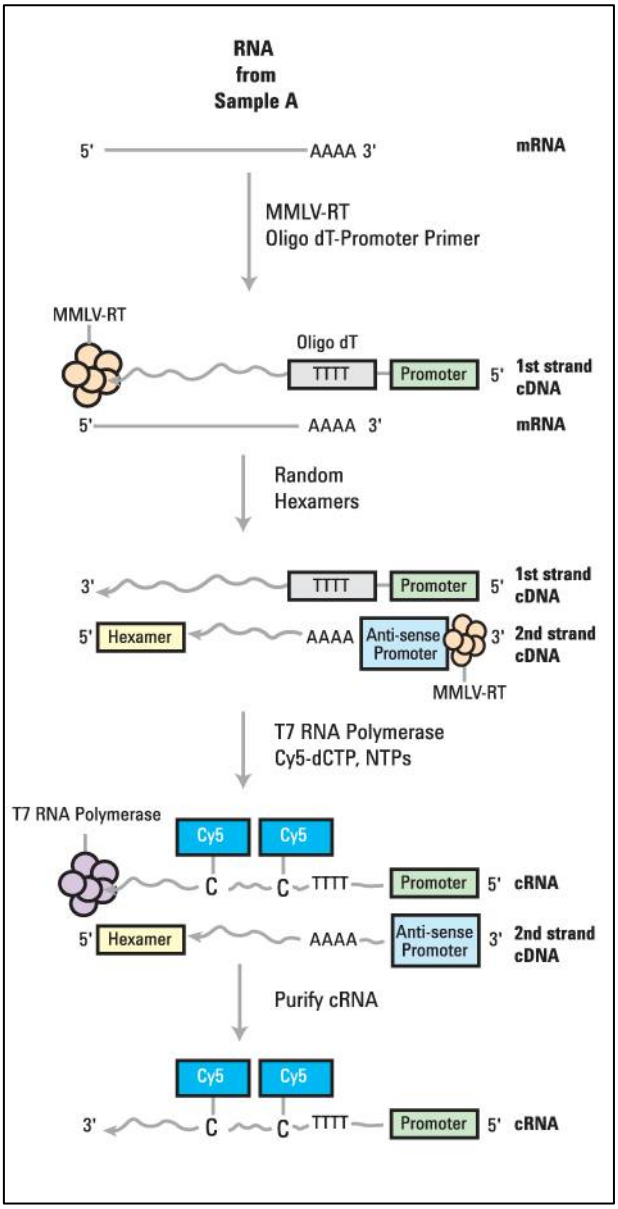



**Figure 2-3.** Procedural overview of the linear amplification labelling step.

Agilent Technologies Fluorescent Linear Amplification kit (Agilent Technologies, USA) was used for the probe labelling from total RNA (Section 2.4.2.1-3). Instead of labelled cDNA probes, this kit produces labelled cRNA probes from Cy3/5-CTP (Perkin Elmer, USA) as the end product, with an amplification of RNA from a smaller

amount of starting material. Only 5 µg of total RNA was required, compared to 25 µg above.

Five µl of T7 promoter primer was added to 5 µg of total RNA. This was topped up with nuclease-free water (Gibco BRL, USA) to 9.5 µl. The primer and template was denatured at 65°C for 10 mins and chilled on ice. Reverse transcriptase (MMLV-RT) was then added and incubated at 40°C for 4 h, after which it was inactivated at 65°C for 15 min. Four µl of the respective Cy3/5-CTP dyes were added together with the T7 polymerase for the amplification step. The samples were then incubated at 40°C for 1 h. A workflow of this new labelling strategy is shown in Figure 2-3.

Purification of the labelled products was carried out through the precipitation of labelled cRNA. This was achieved through the addition of 80 µl of 4.0 M LiCl and stored at -20°C overnight. The cRNA was pelleted through centrifugation at maximum speed for 20 min. The pellet was then rinsed with 70% ethanol, and the pellet was allowed to dry at room temperature for 10 min. Each pellet was resuspended in 100 µl of nuclease-free water (Gibco BRL, USA). The amount of incorporated CyDyes in the cRNA was again quantified as described in Section 2.4.2.5 and equal amounts were combined. Ten µl of 25x Fragmentation Buffer was added and incubated at 60°C for 30 min in the dark. This prevents self-annealing of the RNA during long hybridization. Just before hybridization, 250 µl of 2x Hybridization Buffer was added to the mixture.

Hybridization was carried out as described in Section 2.4.3, except that the automated Lucidea SlidePro Hybridizer was not used. Manual hybridization was carried out on

the SureHyb Chambers (Agilent Technologies, USA) instead, according to the manufacturer's protocol. The hybridization mixture was dispensed on the microarray and enclosed inside the hybridization chamber. Care was taken to ensure that the bubbles within the chamber were free to move when rotated. The whole chamber was then mounted in a preheated hybridization oven (Labnet ProBlot 6TM, USA) at 60°C for 17 h with a rotational speed of 4 rpm. Washing of the slides was also carried out manually. After hybridization, the microarray was disassembled from the hybridization chamber and immersed in a 250 ml Pyrex beaker filled with Wash Solution 1 (Appendix 3g) under a magnetic stirrer (Stuart Scientific SB162, UK) for 10 min. The microarray was then transferred to a second beaker filled with Wash Solution 2 (Appendix 3h) maintained at 4°C for 5 min. Finally, the microarray was carefully lifted out and a nitrogen-gas gun was used to dry the slide. The microarray slide was then immediately scanned (Section 2.4.4).

### 2.4.6 Data Analysis

Transforming images into gene expression matrix is a non-trivial process (Brazma and Vilo, 2000). The spots corresponding to genes on the microarray should be identified, their boundaries determined, the fluorescence intensity from each spot measured and compared to the background intensity. It should be noted that methods for the analysis of microarray data are still evolving, and there is currently no standard experimental design or method of data analysis for microarray experiments.

#### 2.4.6.1 Image Analysis

The first step in the microarray workflow was to locate the spots within the scanned image. The positions of the spot have to be accurately defined and measurement of fluorescence intensity for each spot was carried out by quantifying the intensities of pixels within (foreground) and outside (background) the spot. GenePix Pro 4.1 (Axon Instruments, USA) was able to automate the spotfinding process by aligning a grid of spots over the scanned image of the array. The file for the grid patterns with the corresponding gene annotations was supplied by Agilent.

#### 2.4.6.2 Quality Control Check

Quality control spots were next checked using the software to ensure that the hybridization process occurred correctly. Negative control spots should not give any signal at all, while positive control spots consisting of housekeeping genes should usually show up, but may vary depending on experimental conditions. Control spots were identified according to the gene array list supplied by Agilent. The negative controls should have a signal-to-noise (SNR) ratio of less than 3 (SNR = [average signal – average background]/standard deviation of background). Any higher SNR than 3 suggests that the data obtained from the experiment may not be accurate.

Bad spots are defined as having less than 55% of their pixels brighter than the median background intensity at both wavelengths. These bad spots are ‘flagged’ and removed from the list.

A lower limit of detection was determined by calculating the standard deviation of the background subtracted by the median intensities of negative control spots. This

standard deviation was then multiplied by 3 to give the value of the lower limit of detection. Spots on the microarray whose pixel intensities fall below this lower limit were removed from the list. These spots were deemed to be below the threshold for reliable pixel intensity detection.

#### 2.4.6.3 Database Generation and Analysis

The data for all remaining spots were exported as .txt files from GenePix Pro as tab-limited format, and read with Excel XP. Normalization and analyses were performed using BRB ArrayTools 3.1 (National Cancer Institute, http://linus.nci.nih.gov/BRB-ArrayTools.html), which encompasses the statistical software R within the Excel framework. The data was then logarithmic (base 2) transformed before Lowess (or intensity-dependent) normalization (Yang *et al.*, 2002b) was carried out. This will reduce variability in the data across slides due to the differences in the signal intensity values between the 2 dyes. A logarithmic scale was used because transforming expression data to log scale removes much of the proportional relationship between random error and signal intensity. Distributions of replicated logged expression values and log ratios tend to be normal (Nadon and Shoemaker, 2002). Since the CyDyes were switched on the replicate microarrays, the calculation of the $Log_2$ ratios had to be switched accordingly. This was performed to prevent CyDye bias during incorporation of dyes into certain genes. The virus sample is always compared against the control sample. Positive Log ratios would therefore, indicate that mRNA from the virus sample was in abundance, whereas negative Log ratios would indicate the reverse. The data was then screened for genes with at least a 2-fold change in expression values.

Software from The Institute of Genomic Research (TIGR) were used for the more advanced analysis carried out in the second part of this study. The 2 programmes used were MIDAS and MeV (Saeed *et al.*, 2003).

MIDAS was used to perform a number of transformations on the data to eliminate questionable or low quality data. It also adjusts the measured intensities to facilitate comparisons, and to select those genes that are significantly differentially expressed. Lowess normalization was first performed after importing the microarray data into the software, so that the raw data can be compared across the 2 channels. A flip-dye consistency checking was then carried out between the flip-dye pairs of experiments. Inconsistent spots whose intensities deviate by more than 2SD are removed from the list. Finally slice analysis was carried out to classify and filter spots based on their expression levels. Differentially regulated genes can be identified at a 95% confidence interval.

To perform the clustering analyses to determine the trends in gene expressions, MeV was used to carry out clustering algorithms and to generate the graphical output. Various clustering methods were carried out to determine the best method. The functions of these different methods are highlighted in Section 1.7.3.2. Data generated by MIDAS was uploaded into MeV before starting the algorithms. Hierarchical clustering, *K*-means clustering, Self Organizing Tree Algorithm (SOTA), Figures of Merit (FOM) test, and *t*-tests were all carried out individually. The graphical outputs were then inspected and the most coherent clusters were selected. The genes in each cluster were subsequently analysed for any significance or relevance to the virus-host infection system being studied.

Up and down regulated genes or clusters of genes were then separately uploaded into *EASE* (Hosack *et al.*, 2003). This software helps to classify the uploaded genes according to their biological themes or ontology groups. Genes from any functional groups which are significantly over-represented or predominate during the virus infection process are identified. For example, if there are 10 genes known to be involved in apoptosis, and 8 genes relating to apoptosis were found to be differentially regulated, *EASE* will identify the apoptosis ontology group to be highly over-represented. If only 1 gene relating to apoptosis was found, this ontology group will not be selected. Biological categories with Fisher's exact test of *P*-values less than 0.05 were selected.

## 2.5 Indirect Immunofluorescence Microscopy

For immunofluorescence microscopy, the cells were cultivated as described in Section 2.1.5. When the monolayers reached confluency of about 70%, the cells were infected with WN(S)V as before (Section 2.2.2). Each coverslip in the 24-well tissue culture plate was infected with 50 µl of virus and incubated for 1 h at 37°C. After 1 h adsorption, the excess innocula were removed before 1.5 ml of DMEM was added. Mock-infected cells using virus diluent was used as controls. The plate was incubated at various time points until it is ready for immunofluorescence microscopy studies.

The antisera used and their sources are described as below:

**Table 2.1: Antibodies and their working dilution used in IFA.**

| Type of antibody | Name | Dilution |
|---|---|---|
| **Primary antibodies** | Mouse monoclonal anti-tubulinα antibody (Amersham Biosciences, UK) | 1:250 |
| | Mouse monoclonal anti-actin antibody (Chemicon International, USA) | 1:250 |
| | Rabbit polyclonal anti- WNV Envelope protein antibody (gift from Vincent Deubel, Pasteur Institute, France). | 1:500 |
| **Secondary antibodies** | Sheep anti-mouse Ig, flourescein isothiocynate (FITC) linked (Amersham Biosciences, UK) | 1:500 |
| | Donkey anti-rabbit Ig, Texas Red™ linked (Amersham Biosciences, UK) | 1:500 |

The infected and control cells were washed twice with cold PBS (Appendix 1c) and then fixed with cold methanol (Merck, Germany) for 10 min at -20°C. This was followed by a wash in cold PBS for 15 min. The cells were then blocked with cold 0.1% BSA (Appendix 4a) in PBS to prevent non-specific attachment of antibodies.

Primary antibodies were diluted as detailed above in Table 2.1. Fifty µl of the diluted antibodies was spotted on parafilm. Coverslips seeded with cells were then inverted over the drop of antibody and incubated for 1 h at 37°C in a humid chamber. After incubation, the excess antibodies were washed off thrice with cold 0.1% BSA in PBS for 5 min each at room temperature. Species-specific secondary antibodies were appropriately diluted in PBS as detailed in Table 2.1. Coverslips were similarly exposed to the secondary antibodies as above. After incubation, the coverslips were washed three times with cold PBS for 5 min each. For double immuno-labelling, secondary antibodies of different species specificity were added in succession.

A single drop of mountant (Appendix 4b) was placed on ethanol-cleaned glass slides and the coverslips inverted over the mountant. Excess mountant was blotted with lint-free paper (Kimwipe, Kimberly Clark, Canada). Fluorescence was visualised under optical immunofluorescence microscopy (IX81, Olympus, Japan) using oil immersion objectives. FITC (480 nm) and Texas Red (543 nm) emit green and red fluorescence, respectively. When the two fluorescent stains are co-localized, yellow fluorescence would be detected. Where relevant, quantification of the fluorescent intensity was performed using the MetaMorph software (Universal Imaging Corporation, USA).

## 2.6 Quantitative Real-Time PCR

### 2.6.1 List of Oligonucleotides Synthesised During the Project

Appendix 5 lists the oligonucleotides (Proligo, Singapore) that were synthesised and used in the quantitative analyses of RNA transcripts. (+) denotes sense orientation; (-) anti-sense orientation. Primers against target genes for qRT-PCR should optimally be in the range of 100-180 bp to allow for efficient thermal cycling and binding of the SYBR dye. As such, designing a primer based on the 60 bp oligonucleotide probes found on the microarray is insufficient. Sequences for the primers against target genes were primarily sourced from the online database 'PrimerBank' [(http://pga.mgh.harvard.edu/primerbank/) (Wang and Seed, 2003)]. PrimerBank is a public resource for PCR primers, which is designed for gene expression detection or quantification (real-time PCR). If primers for a particular gene of interest are not found in PrimerBank, then 'Primer3' [(http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) (Rozen and Skaletsky, 2000)] was used as a tool to generate optimal primers based on target gene sequences.

### 2.6.2 Real-Time PCR

New samples of total cellular RNA, different from the microarray batch, were extracted as described in Section 2.4.2.1. An additional DNase treatment step was included to remove all contaminating genomic DNA as qRT-PCR is a very sensitive quantification method. The qRT-PCR was performed in two steps; reverse transcription to cDNA and then followed by real-time PCR. For cDNA synthesis, 5 µg of total RNA was reverse transcribed using 200U SuperScript III (Invitrogen, USA) in a total volume of 20 µl containing 500 µM dNTP mix, 5 mM $MgCl_2$, 20 mM DTT, 40U RNaseOUT, primed with 2.5 µM oligo(dT). Reverse transcription was performed at 50°C for 50 min, followed by 85°C for 5 min, according to the manufacturer's protocol.

For the quantification of viral transcripts, the primer used was targeted against the WNV envelope (*E*) gene. Even though a similar SYBR green-based qPCR assay has been reportedly carried out (Papin *et al.*, 2004), this was however on the *E* region of the WNV/NY99 strain. Due to strain variability, a different primer pair for the *E* gene was developed in this study. During reverse transcription, the WN(S)V E (-) primer (Appendix 5) was used, instead of oligo(dT), to target the *E* region specifically during the cDNA synthesis reaction. All the cDNAs were subsequently diluted 1:10 with sterile Nanopure $H_2O$.

For real-time PCR, 25 µl reaction mixture containing 2 µl of diluted cDNA, 12.5 µl of Platinum SYBR Green (Invitrogen, USA) and 0.2 µM of both forward and reverse primers (Proligo, USA) (Appendix 5) was used. A negative template control that contained all SYBR green reagents except DNA was performed in parallel. Reactions

were cycled at 50°C for 2 min and then 95°C for 2 min, followed by 45 cycles of 95°C for 15 s, 60°C for 30 s and 72°C for 30 s, followed by a melting curve analysis. These were performed on the iCycler iQ (Bio-Rad, USA). Each gene was quantified at least 3 times, with a triplicate sample each time. This was to increase the statistical power and to average the readings.

A calibration curve containing 5 points ranging from 100 fg to 1 ng of cDNA was used as a standard. The hypoxanthine guanine phosphoribosyltransferase (HPRT1) gene was used as an internal control for normalization (Johnson *et al.*, 2004; Vandesompele *et al.*, 2002), as it is a putative housekeeping gene. Other common housekeeping genes, such as G6PH and GAPDH, were found to be differentially expressed during virus infection. The threshold cycle ($C_T$) values were then translated into relative copy numbers of cDNA by using the $2^{-\Delta\Delta C}{}_T$ method of calculation (Livak and Schmittgen, 2001) as follows:

$$\text{Relative change} = 2^{-\Delta\Delta C}{}_T$$

$$\text{where } \Delta\Delta C_T = (C_{T,\,\textit{Target}} - C_{T,\,\textit{HPRT1}})_{\text{virus}} - (C_{T,\,\textit{Target}} - C_{T,\,\textit{HPRT1}})_{\text{control}}$$

# 3.0 RESULTS – Comparison between HeLa and A172 cells

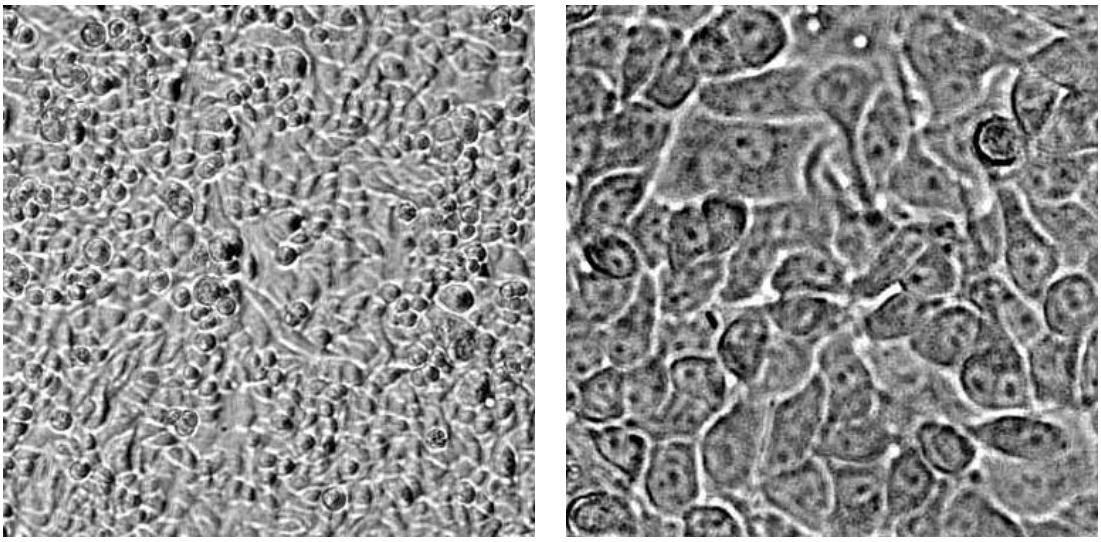
Numerous studies on WNV are carried out on non-human host cells (e.g. Vero, C6/36, etc.) *in vitro*, but few are done on human cell lines and thus, the virus-host interactions have not been clearly established hitherto. In this study, West Nile (Sarafend) [WN(S)] virus was used to infect human cells. HeLa (human cervical carcinoma) cells were initially used in this study as it was a readily available cell line in laboratories and is of human origins. Studies on infectivity rates were performed on HeLa cells, but were found to be relatively poor. A172 (human brain glioblastoma) cells were subsequently obtained and were found to be highly susceptibility to WN(S) virus infection. A global genomics method was therefore carried out to compare the infectomics between HeLa and A172 cells.

## 3.1 West Nile (Sarafend) Virus [WN(S)V] Infection on HeLa Cells

When WN(S)V was used to infect HeLa cells, it was found that cytopathic effects (CPE) were not extensive when compared to Vero cells, from which the virus stock was obtained. HeLa cells generally remained quite intact after initial infection and managed to continue propagating slowly in the maintenance medium (2% FCS).

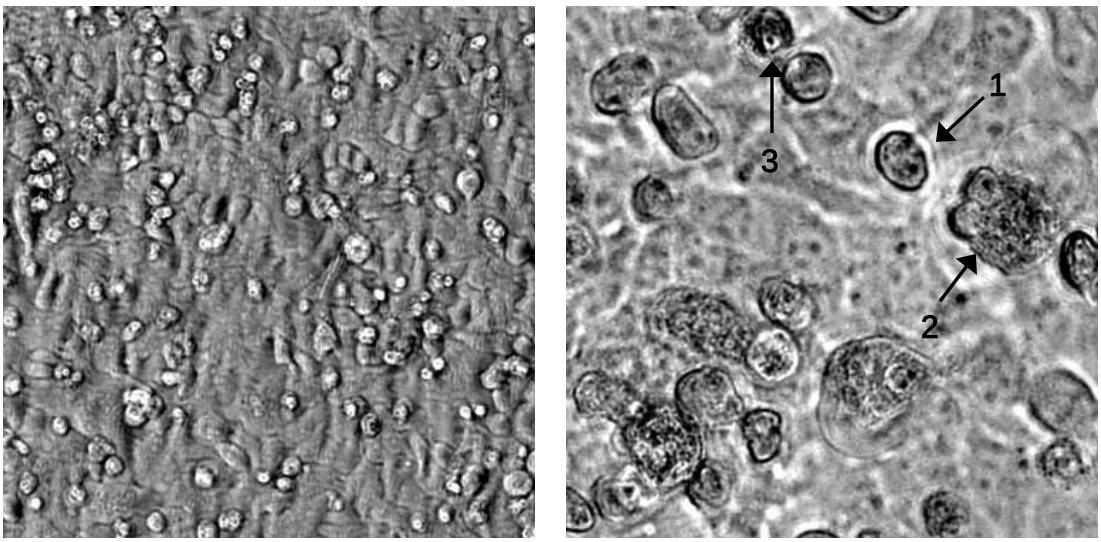
Figure 3-1 shows mock-infected (control) HeLa cells and Figure 3-2 shows WN(S) virus-infected HeLa cells. It was observed that HeLa formed a packed monolayer typical of epidermal cells. Virus-infected HeLa cells showed some signs of CPE with some cell-rounding and cell lysis. Loss of plasma membrane integrity can be seen in

some of the infected cells. Cell lysis or lifting of cells from the growing substrate were, however, less prominent when compared to A172 cells.



[A] [B]

**Figure 3-1.** Normal epidermal cell morphology is observed for the mock-infected control HeLa cells when viewed under phase-contrast microscopy. The magnification is 100x and 400x in [A] and [B], respectively.



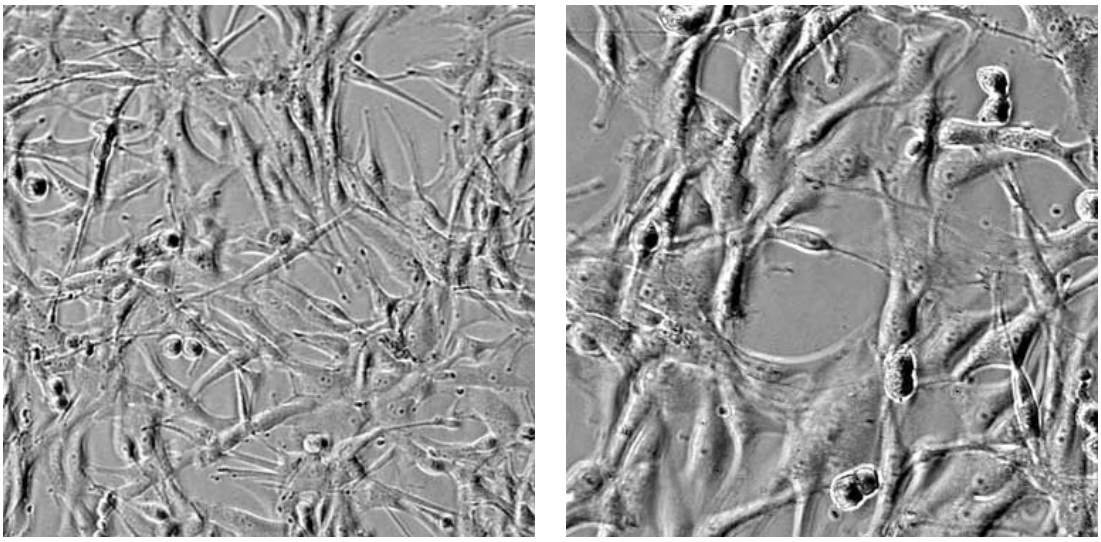

[A] [B]

**Figure 3-2.** Prominent cell-rounding (arrow-1) is observed in the WN(S) virus-infected HeLa cells. Some of the cells exhibit a loss of plasma membrane integrity or membrane blebbing (arrow-2), while others exhibit highly condensed nuclear material (arrow-3). The HeLa cells generally exhibit less severe CPE when compared to A172 cells (Figure 3-4). The magnification is 100x and 400x in [A] and [B], respectively.

## 3.2 West Nile (Sarafend) Virus Infection on A172 Cells

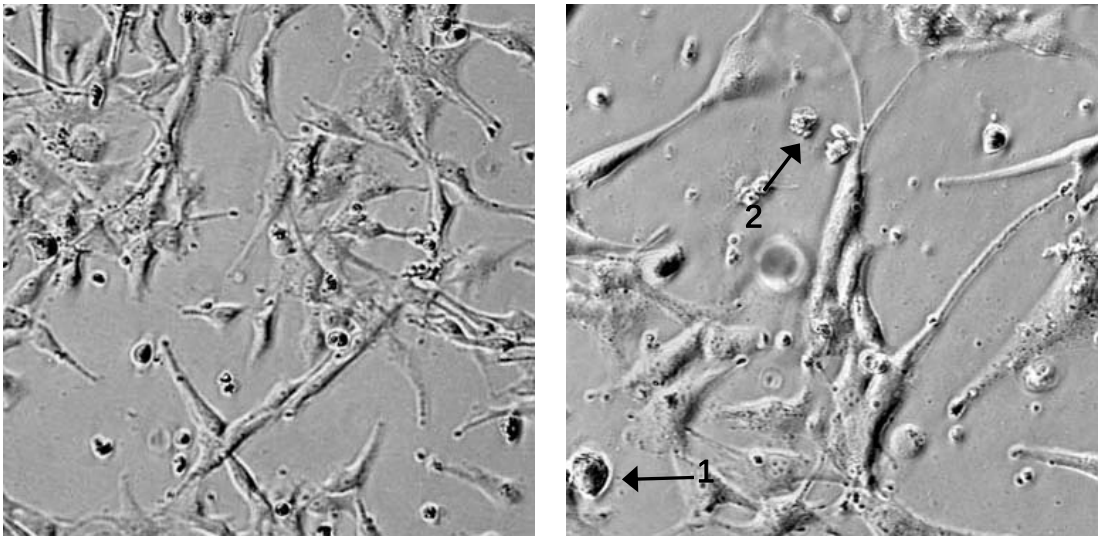
As WNV is known to cause encephalitis and targets the brainstem, WN(S) virus was therefore expected to show enhanced cytopathic effects on A172 cells, which is a brain glial blastoma cell line.

Figure 3-3 shows mock-infected (control) A172 cells and Figure 3-4 shows WN(S) virus-infected A172 cells, where widespread CPE was found. Extensive lifting of infected cells and lysis of cells were observed in the virus-infected cells. Cell shrinkage and condensation were also readily observed, with many fragmented cells. On the whole, A172 cells were found to exhibit more severe CPE compared to HeLa cells.



[A] [B]

**Figure 3-3.** Normal neuro-glial morphology is observed for the mock-infected A172 cells when viewed under phase-contrast microscopy. The magnification is 100x and 400x in [A] and [B], respectively.




[A] [B]

**Figure 3-4.** Extensive CPE with cell-lifting off the growing substrate are observed for WN(S)V-infected A172 cells. Cell shrinkage and condensation are observed (arrow-1). Most of the cells have lysed due to the aggressive virus infection (arrow-2). A172 cells exhibit much more severe CPE compared to HeLa cells (Figure 3-2). The magnification is 100x and 400x in [A] and [B] respectively.

## 3.3 Plaque Assay Studies

The virus stock for this study was produced in Vero cells, which was the preferred cell line for growing WN(S)V with high titres of $10^8$ PFU/ml. The virus was not adapted by passaging through the various cell lines to provide a basal level for comparison.

The virus titre obtained from WN(S)V-infected HeLa cells was consistently around $10^6$ PFU/ml even after nearly 48 h of infection. In comparison, the virus titre obtained from WN(S)V-infected A172 cells was in the range of $10^8$ PFU/ml after just 24 h of infection. The A172 cells were therefore more permissive to WN(S)V infection compared to HeLa cells. The virus titres in the HeLa cells were about 100-fold lower than that in infected-A172 cells. Figure 3-5 shows the histogram from the plaque assay results.

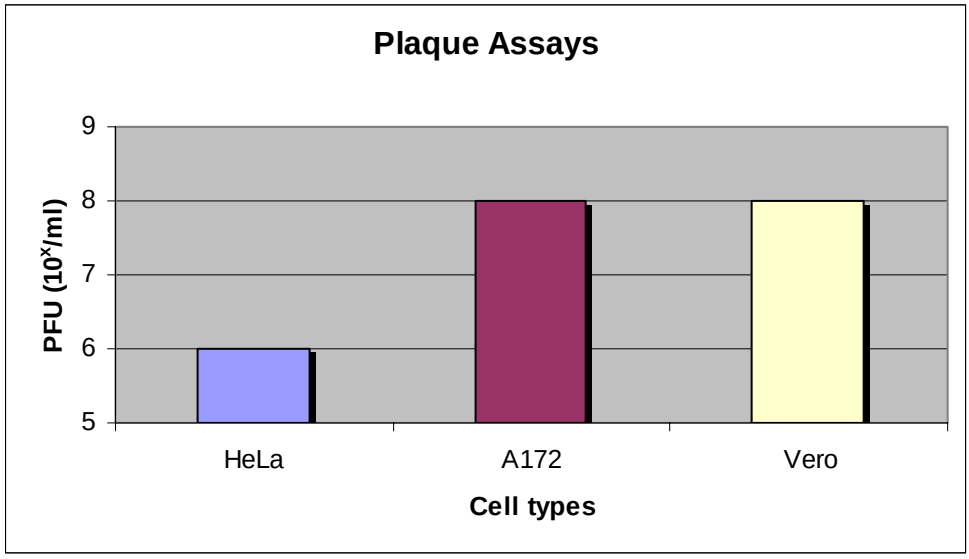



**Figure 3-5.** WN(S)V titres obtained from three different cell lines.

## 3.4 Quantitative Real-Time PCR (qPCR)

Another method was used to determine the efficiency of virus replication in the two different cell lines. qPCR is reported to be the most sensitive method to quantify the number of copies of viral genomic RNA within the cell, and has been used to quantify the virus yield in clinical settings. In this study, the target of the qPCR was against the WN(S)V envelope (E) gene, which forms the envelope protein of the virus. The procedure for this was described in Section 2.6.2.

Figure 3-6 shows the standard curve obtained from a serial dilution of the WN(S)V E gene. The correlation coefficient of the standard curve has a value of $R^2 = 0.998$, thus representing a closely fit curve encompassing the 6-fold dilution range. It can therefore be inferred that test results which fall within this range will be highly accurate. From the slope of the curve, a 10-fold dilution represented a $C_T$ change of 4.07. Figure 3-7 shows the evenly-spaced amplification plot obtained for the dilution series. Figure 3-8 shows the amplification plot obtained for the WN(S)V E gene qPCR tests on A172 and HeLa cells from one set of experiment, and the corresponding dissociation curve (denoting specificity) is shown in Figure 3-9.

Tables 3-1 and 3-2 give the data obtained for the qRT-PCR studies for the relative quantity of the E gene transcripts at 12 h and 24 h p.i., respectively. From the analysis, the amount of E gene transcripts are 8.18x and 4.48x more in A172 than HeLa cells, at 12 h and 24 h p.i., respectively.

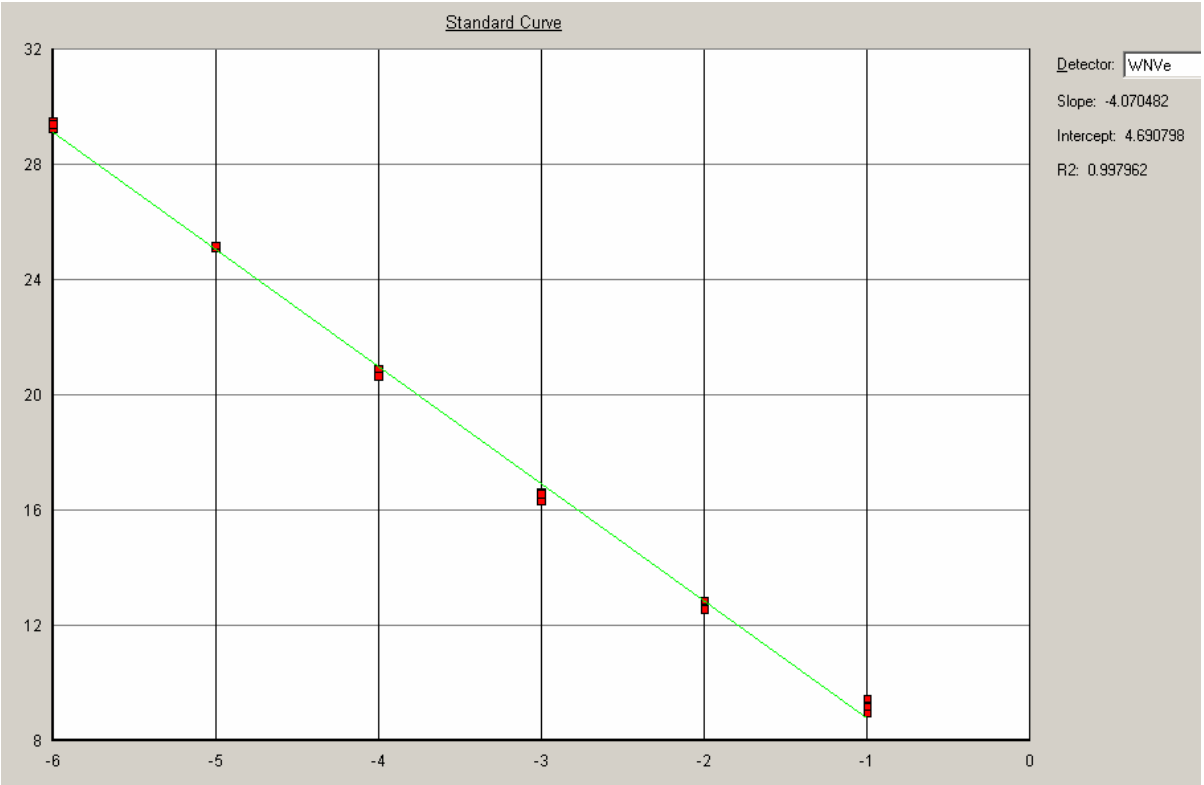



**Figure 3-6.** Standard curve for WN(S)V E gene.

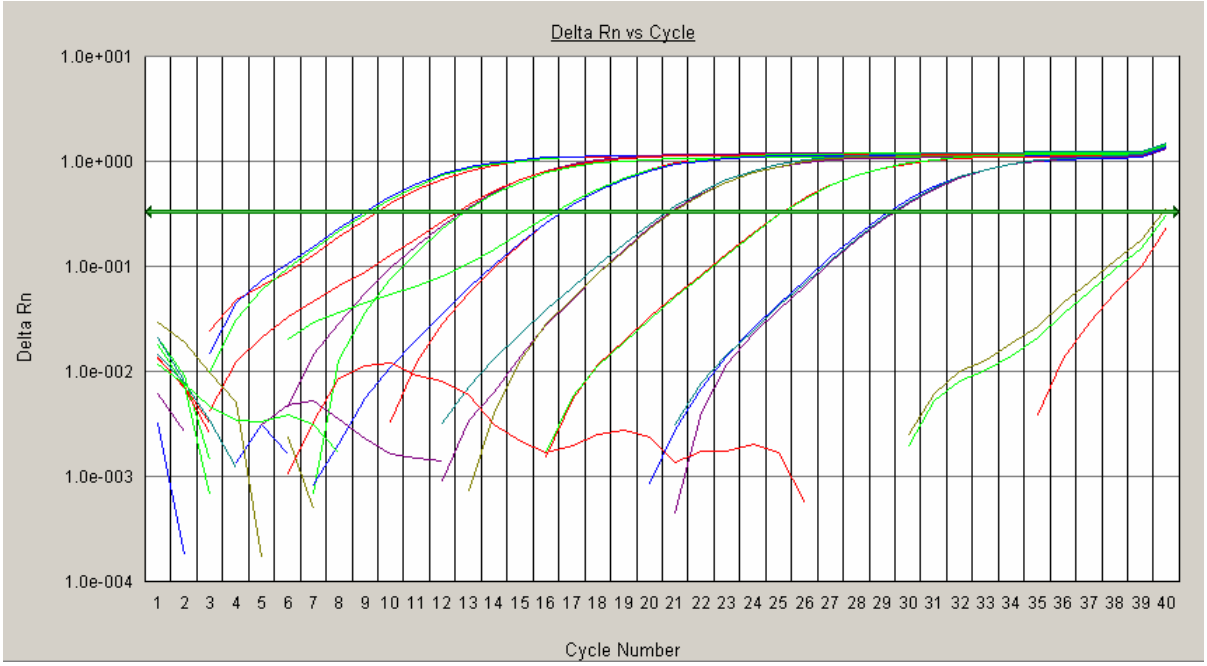



**Figure 3-7.** Amplification plot for dilution series of WN(S)V E gene target. The samples were ten-fold serially diluted, and were carried out in triplicates. Each line denotes one sample.

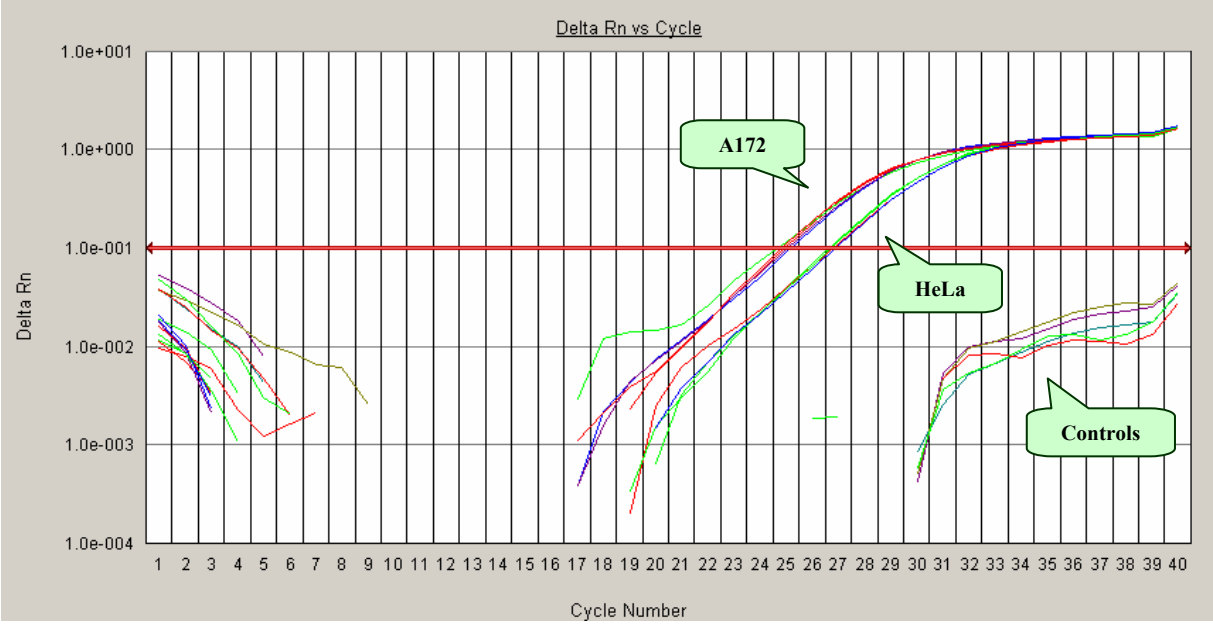



**Figure 3-8.** Amplification plot for WN(S)V E gene in A172 and HeLa cells at 24 h p.i.

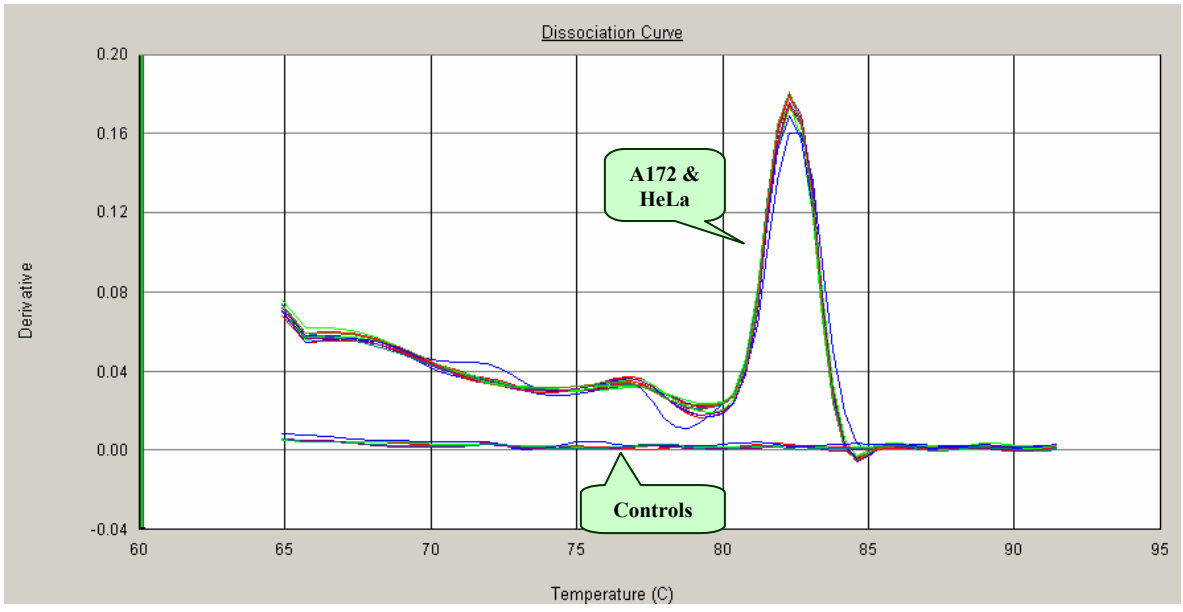



**Figure 3-9.** Dissociation (melt) curve for qRT-PCR.

**Table 3-1.** Data from qRT-PCR on WN(S)V E gene at 12 hours p.i.

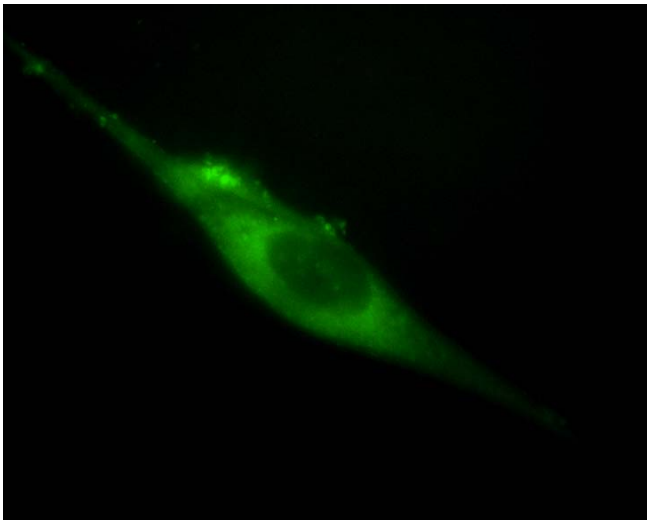
| Test number | $C_T$ (HeLa) | $C_T$ (A172) | $\Delta C_T$ | Fold Change |
|---|---|---|---|---|
| 1 | 36.42 | 32.95 | 3.47 | 8.52 |
| 2 | 31.03 | 26.89 | 4.14 | 10.17 |
| 3 | 29.48 | 27.10 | 2.38 | 5.84 |
| **Average** | | | | **8.18** |

**Table 3-2.** Data from qRT-PCR on WN(S)V E gene at 24 hours p.i.

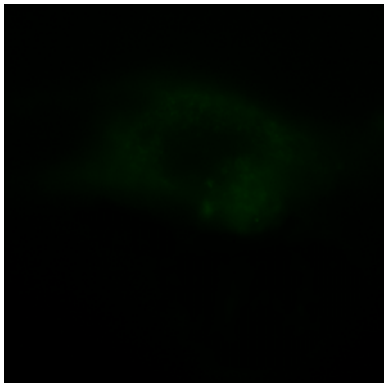
| Test number | $C_T$ (HeLa) | $C_T$ (A172) | $\Delta C_T$ | Fold Change |
|---|---|---|---|---|
| 1 | 26.71 | 24.92 | 1.79 | 4.40 |
| 2 | 35.14 | 32.95 | 2.19 | 5.38 |
| 3 | 43 | 41.51 | 1.49 | 3.66 |
| **Average** | | | | **4.48** |

## 3.5 Immunofluorescence Microscopy of West Nile (Sarafend) Virus

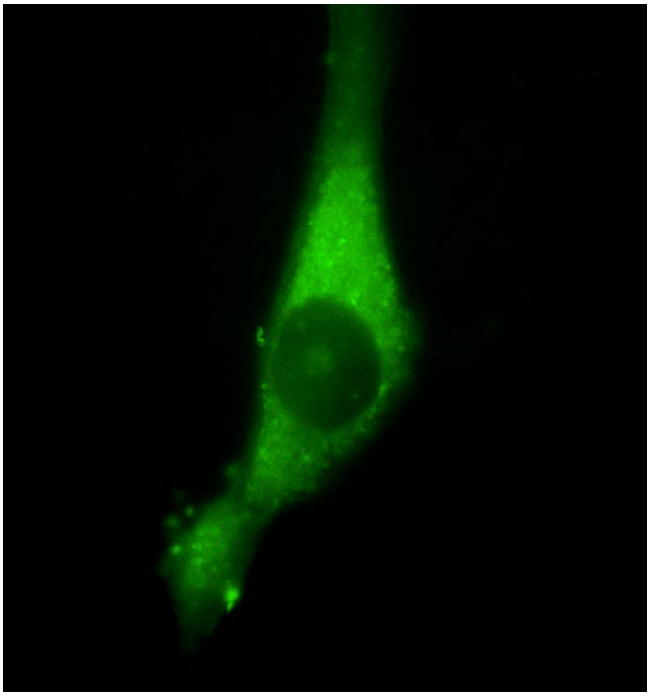
Immunofluorescence (IF) microscopy was carried out to study if the distribution of virus proteins within the cell was the same between the two cell lines, in an attempt to understand the differences in the permissiveness of the two cell lines. Figures 3-10 and 3-11 show the fluorescence images of the cells at 24 h p.i. It was observed that the fluorescence was much more intense in WN(S)V-infected A172 cells (Figure 3-10) than in HeLa cells (Figure 3-11).



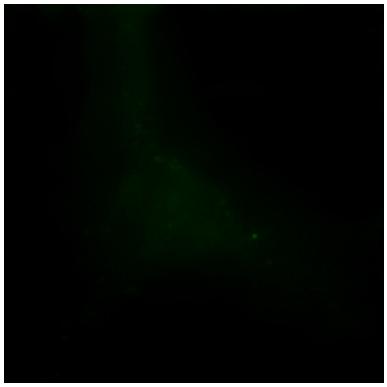
[A]



[C]

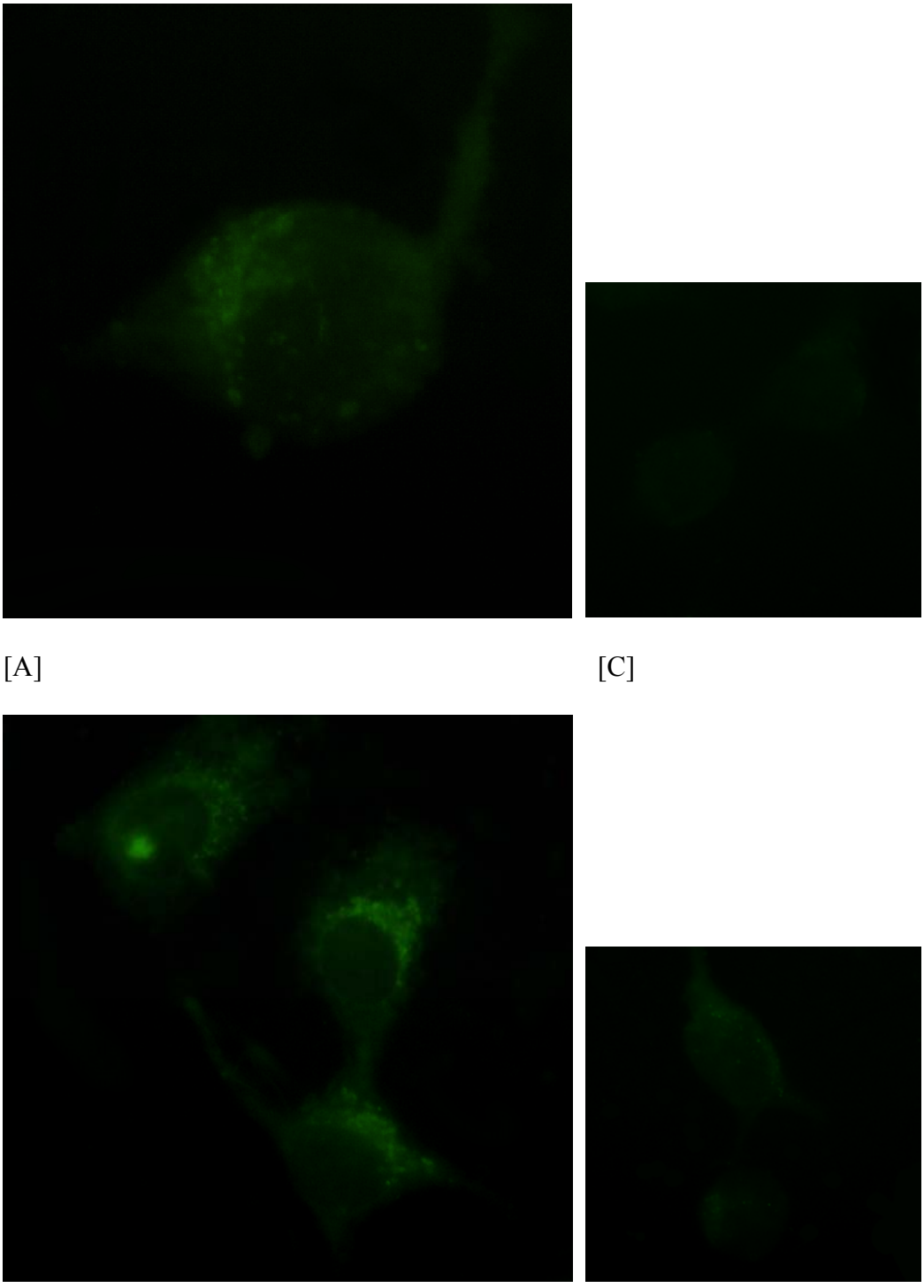


[B]



[D]

**Figure 3-10.** [A] and [B] show WN(S)V-infected A172 cells at 24 h p.i., whereas the [C] and [D] shows mock-infected control A172 cells. The primary antibody is the anti-E protein antibody. Fluorescence staining is detected more strongly at the perinuclear region of the infected cell. Some bright speckles are seen at the plasma membrane. This represents the maturing virus particles. A slight polarization of the staining can be observed in the infected cell. [C] and [D] show negligible fluorescence in the mock-infected cells.



[A] [C]

[B] [D]

**Figure 3-11.** [A] and [B] show WN(S)V-infected HeLa cells at 24 h p.i., [C] and [D] show mock-infected control HeLa cells. The primary antibody is the anti-E protein antibody. Fluorescence staining is detected mainly at the perinuclear region of the infected cell, but is less dispersed when compared to the A172 cells (Figure 3-10). Some speckles and polarized staining are also observed. The fluorescence intensity is much lesser compared to A172 cells (Figure 3-10). [C] and [D] show negligible fluorescence in the mock-infected cells.

As the MetaMorph software that was used to capture the IF images was able to quantify the amount of fluorescence in the images, the intensity of the fluorescence was measured in the two infected cell lines. Table 3-3 presents the data from this finding. From the intensity readings, A172 cells show 3.98x greater fluorescence compared to HeLa cells. This value corresponds closely to the value obtained from qPCR.

**Table 3-3.** Intensity of fluorescence within infected host cells

| **Cell** | **HeLa** | **A172** |
|---|---|---|
| 1 | 16.48 | 62.61 |
| 2 | 16.07 | 65.28 |
| 3 | 18.65 | 64.21 |
| 4 | 15.71 | 67.04 |
| 5 | 15.35 | 68.33 |
| 6 | 16.95 | 64.53 |
| 7 | 17.05 | 67.24 |
| 8 | 15.34 | 64.68 |
| **Average** | **16.45** | **65.49** |
| **Fold Change** | **3.98x** | |

## 3.6 Global Genomics Studies on HeLa and A172 Cells

As different rates of virus infection and replication was observed between the two cell lines, a global transcriptomic study was carried out to determine the host factors that may contribute to determine the efficacy of the virus infection in different hosts.

### 3.6.1 Total RNA Isolation

Total RNA had to be first extracted from the cells in order to carry out microarray experiments for transcriptomic studies. As RNase is ubiquitous anywhere, extra precautions are necessary when handling RNA. Total RNA isolation was carried out as described in Section 2.4.2.1.

Microarray experiments were initially conducted according to the methods described in Sections 2.4.2 – 2.4.3. Essentially, these methods utilized Amersham's probe labelling kits and automated hybridization workstation. Two sets of experiments were carried out for each cell line. Each set consisted of a virus-infected and mock-infected control samples. This was carried out in duplicates, but with a dye-swap between the virus-infected and control samples.

The quantity and purity of eluted total RNA are shown in Table 3-4. The purity levels were in the range of 1.6 – 2.1 before proceeding with the labelling step. It was found that the yield of total RNA was markedly lower in the WNV-infected A712 cell samples, compared to the rest of the samples. The explanation for this could be that the cytopathic effects in infected A172 cells were more advanced compared to

infected HeLa cells (Section 3.2). Nevertheless, the amount of RNA harvest from each batch was sufficient to proceed with the probe labelling.

**Table 3-4:** Quantity and purity of RNA samples.

| Sample | | Dye | $A_{260}$ | $A_{280}$ | Concentration (µg/µl) | Yield (µg) | Purity |
|---|---|---|---|---|---|---|---|
| **HWN1** | **HWNV1** | Cy3 | 0.605 | 0.367 | 1.21 | 72.6 | 1.65 |
| | **HWNC1** | Cy5 | 0.811 | 0.447 | 1.62 | 97.3 | 1.81 |
| **HWN2** | **HWNV2** | Cy5 | 0.336 | 0.206 | 1.12 | 67.2 | 1.63 |
| | **HWNC2** | Cy3 | 0.260 | 0.151 | 0.867 | 52.0 | 1.72 |
| **AWN1** | **AWNV1** | Cy3 | 0.436 | 0.237 | 0.496 | 29.76 | 1.97 |
| | **AWNC1** | Cy5 | 0.484 | 0.269 | 1.312 | 78.72 | 1.86 |
| **AWN2** | **AWNV2** | Cy5 | 0.481 | 0.245 | 0.450 | 27.00 | 1.84 |
| | **AWNC2** | Cy3 | 0.498 | 0.285 | 1.870 | 112.2 | 1.90 |

**Sample:** H – HeLa, A – A172, WNV – Virus infected sample, WNC – Control sample.

**$A_{260}$/$A_{280}$:** UV Absorbance readings.

**Concentration** = $A_{260}$ X 40 X dilution factor (50)

**Yield** = Concentration X volume (60)

**Purity** = $A_{260}/A_{280}$

### 3.6.2 Integrity of Isolated Total RNA

After the extracted samples had been quantified, it was crucial for the integrity of RNA to be determined. Electrophoresis of samples on denaturing formaldehyde-agarose (FA) gel will elucidate if the RNA samples had degraded. Figure 3-12 shows a representative FA gel image from the HeLa cell samples. Intact total RNA samples

will give two sharp ribosomal bands for 18S (1.9kb) and 28S (5.0kb) rRNA (Figure 3-12). mRNA appeared as a very faint smear in between these two bands. Degraded RNA samples will give indistinct small fragments that will produce inconsistencies in cDNA synthesis and thus, generate inaccurate microarray results. After the RNA integrity had been verified to be intact, the RNA samples were then reverse transcribed, labelled, and purified, as described in Sections 2.4.2.4. The CyDye used are denoted in Table 3-4.

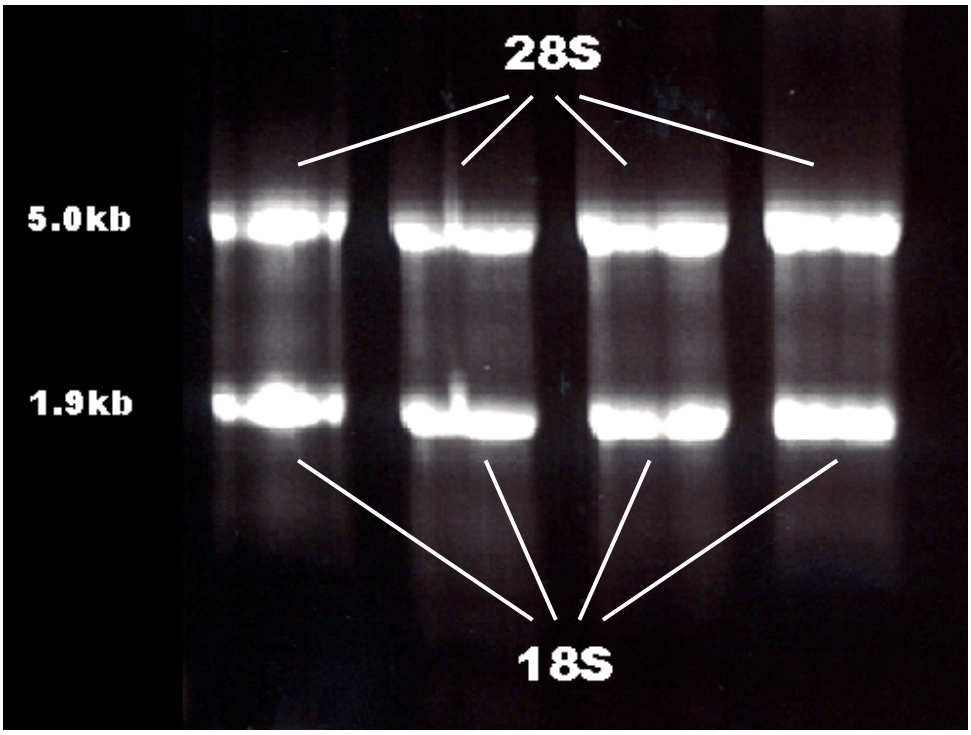



**Figure 3-12.** Diagram shows the intact ribosomal 28S and 18S RNA bands from HeLa cell samples. The faint smear represents mRNA. 3µg of total RNA was loaded per lane. L-R: HWNV1, HWNC1, HWNV2, HWNC2.

### 3.6.3 Quantification of Incorporated Fluorescent Nucleotides

After the RNA had been reverse transcribed into labelled cDNA probes, it was crucial to quantify the amount of incorporated fluorescent nucleotides, and combine equal amounts of the labelled probe for microarray hybridization. The procedure for fluorescent nucleotide quantification is described in Section 2.4.2.5. Figure 3-13

shows the general experimental strategy for RNA labelling. Briefly, total RNA was reverse transcribed to produce labelled cDNA. The RNA template was subsequently degraded. The labelled cDNA from control and virus-infected samples were combined in equal amounts and allowed to hybridize onto a microarray. Table 3-5 shows the quantity of purified cDNA probes harvested by UV spectrometry.

**Table 3-5:** Quantity of incorporated fluorescent nucleotides.

| Sample | | Cy3 ($A_{550}$) | Cy5 ($A_{650}$) | Quantity of CyDye (pmol) |
|---|---|---|---|---|
| H24WN1 | **H24WNV1** | 0.030 | --- | 24.0 |
| | **H24WNC1** | --- | 0.041 | 19.7 |
| H24WN2 | **H24WNV2** | --- | 0.049 | 23.5 |
| | **H24WNC2** | 0.031 | --- | 24.8 |
| A24WN1 | **A24WNV1** | 0.032 | --- | 25.6 |
| | **A24WNC1** | --- | 0.045 | 19.2 |
| A24WN2 | **A24WNV2** | --- | 0.040 | 21.6 |
| | **A24WNC2** | 0.034 | --- | 27.24 |

**Sample:** H – HeLa, A – A172, WNV – Virus infected sample, WNC – Control sample.

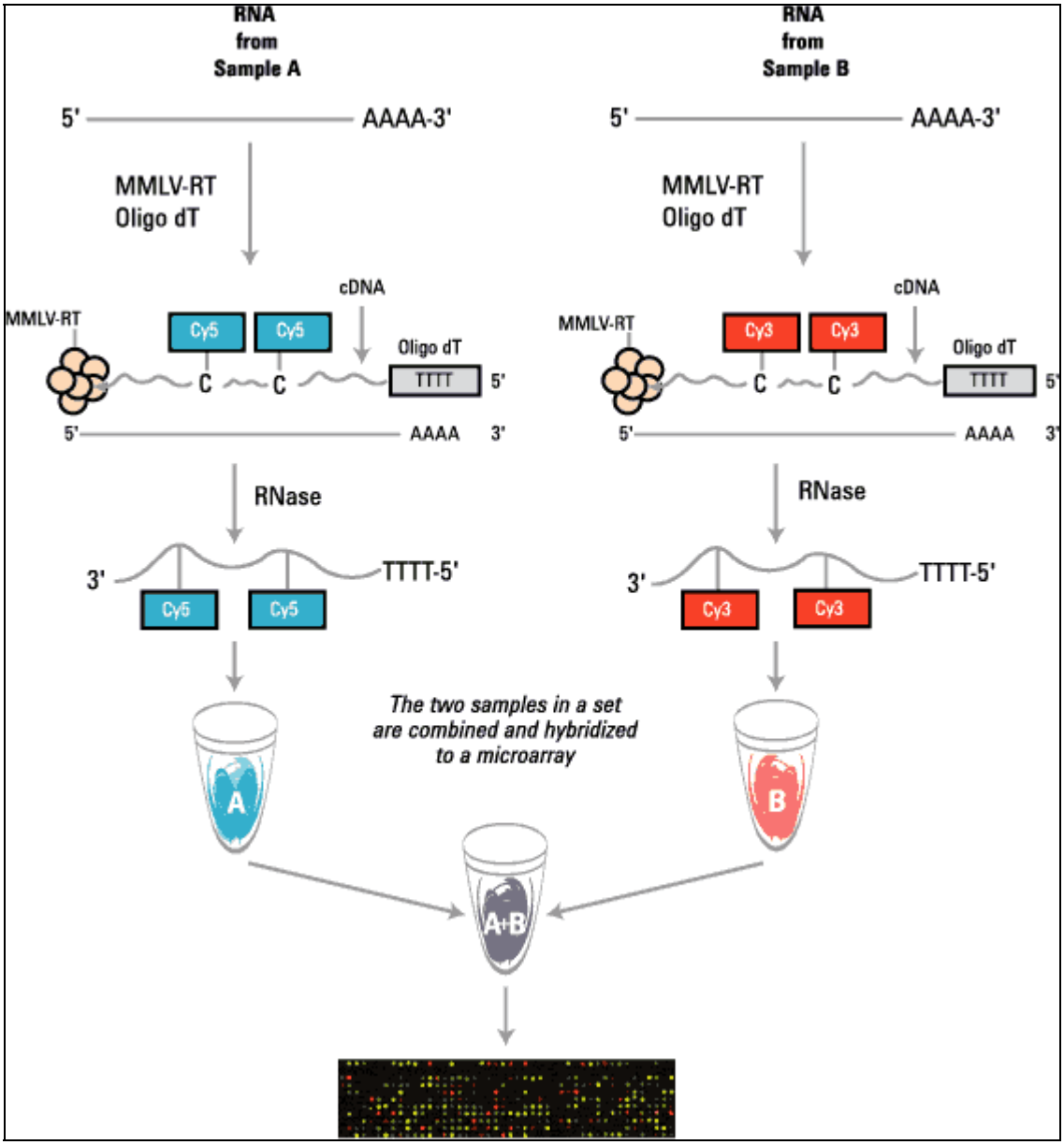



**Figure 3-13.** RNA labelling strategy. Total RNA from control and virus-infected samples are separately reverse transcribed with MMLV-RT using oligo-dT as primers. This is carried out in the presence of CyDye-labelled nucleotides (dCTP). The RNA template is then degraded with RNase, leaving the labelled cDNA probes, which have to be purified from unlabelled probes. Equal amounts of labelled cDNA probes are combined and allow to hybridize onto a microarray. The CyDyes used are switched across different sets of experiments as shown in Table 3-4. This is to prevent CyDye bias.

### 3.6.4 Microarray Images

After microarray hybridization (on Amersham's Lucidea SlidePro) and scanning were carried out, the image file was analyzed. Figure 3-14 shows the raw images generated from the scanner for two of the microarray slides.

Initial microarray scans displayed a 'blotchy' pattern on the first few microarrays. This could be due to the formation of bubbles during the hybridization step, which may interfere with the probe binding to the target. Spots on these blotches were removed if they do not fulfil the quality control tests as highlighted in Section 2.4.6.2. Such blotchy patches will affect the data quality generated, and many of these spots were removed from the final analysis, resulting in a lower number of genes that appeared in the final list. After several attempts, a 'clean' pattern was finally obtained, and is shown in Figure 3-14B, and these were used for downstream data analyses.

Figure 3-15 shows the landmark spots used for image orientation. The number of spots found at the corners marks the orientation positions of the slide. These spots can be correlated with Figure 3-16, which shows the distribution of negative and positive control spots on the microarray slides. Spots marked with (+)Pro25G denote positive-control spots, (-)3xSLv1 denote negative-control spots, while QC spots are for the manufacturer's own quality control of each slide.

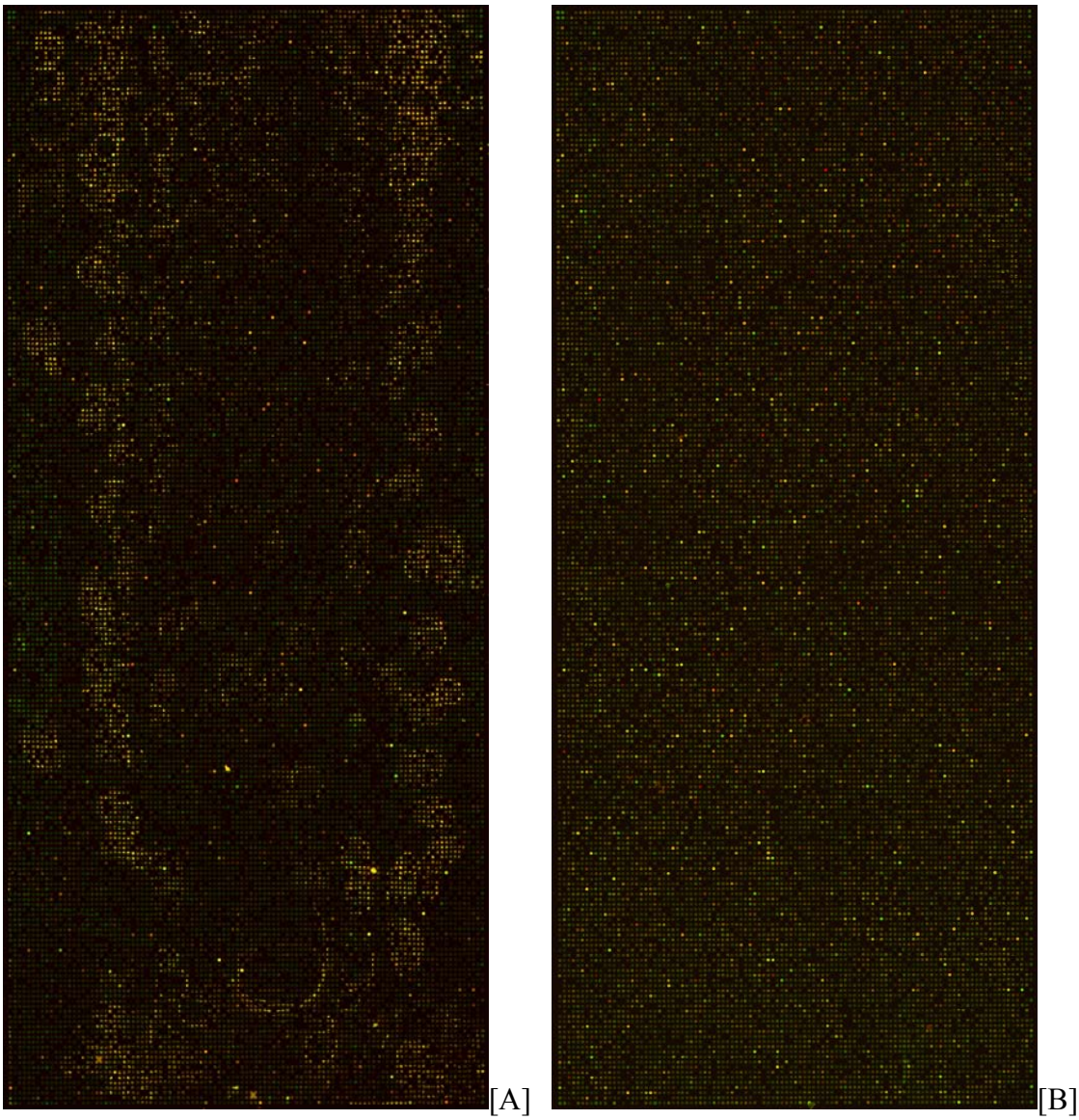



**Figure 3-14.** Raw scans of HWN1 [A] and AWN2 [B]. Initial microarray scans [A] displayed a 'blotchy' pattern on the microarray. This could be due to the formation of bubbles during the hybridization step. AWN2 [B] showed the best hybridization pattern with an absence of any patches.

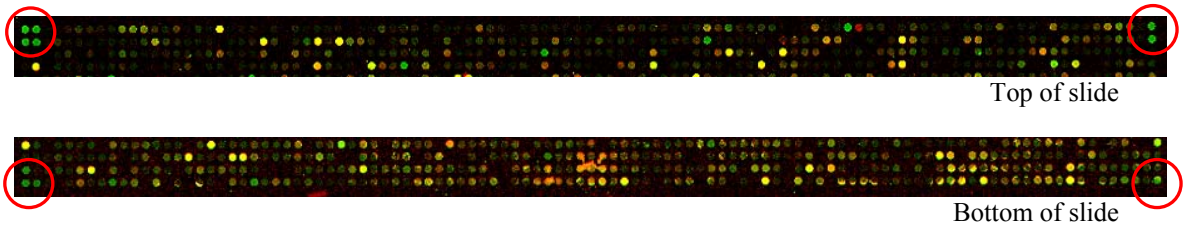
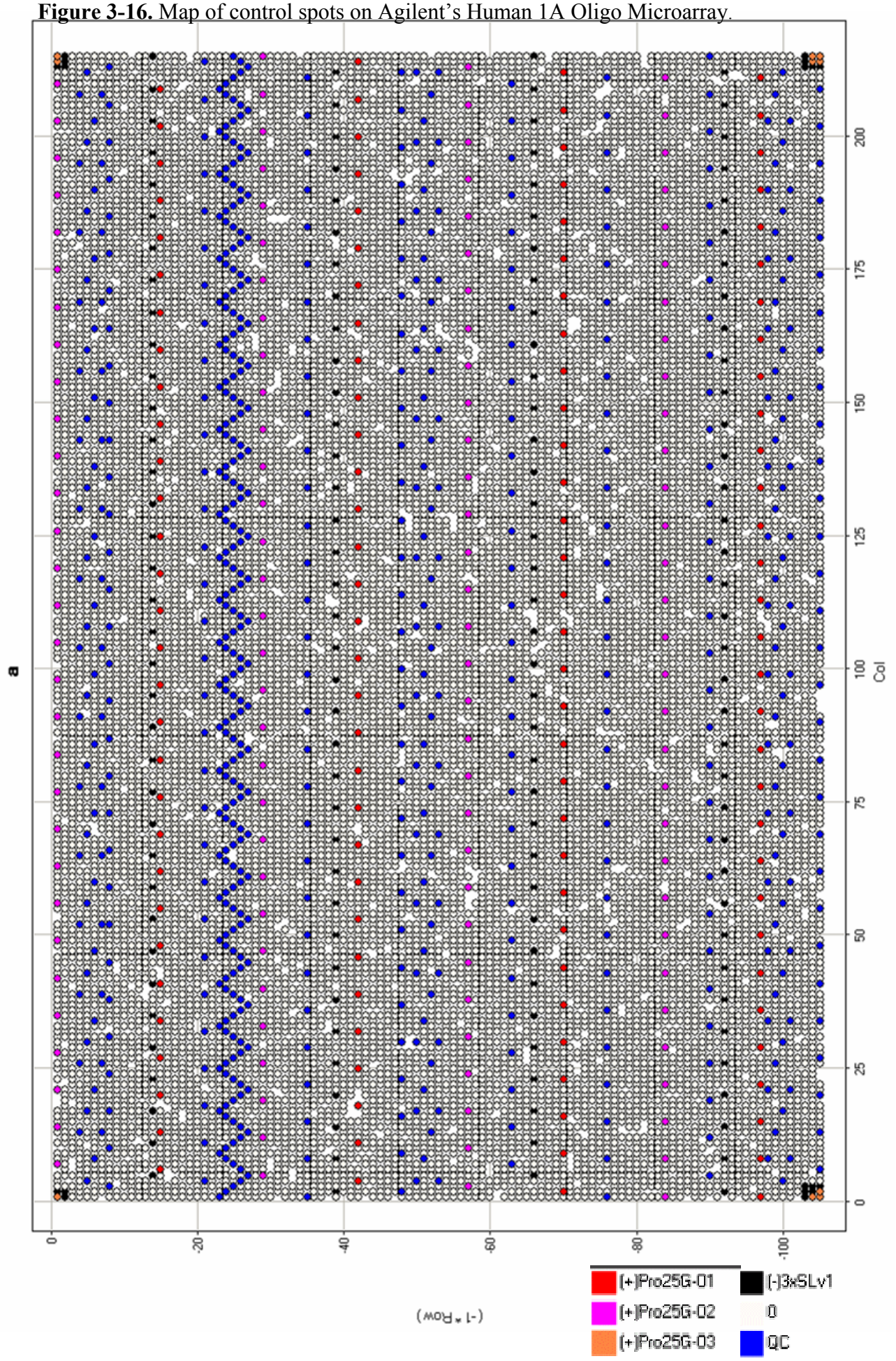



**Figure 3-15.** Landmark spots for slide orientation are circled. The number of green spots at the corners marks its orientation. The above sections were taken from HWN2.

**Figure 3-16.** Map of control spots on Agilent's Human 1A Oligo Microarray.



### 3.6.5 Microarray Image Analysis

The positions of the spots on the images were determined by placing a grid for spot locations over the image, as shown in Figure 3-17. GenePix Pro was then able to automatically fit every single grid spot over the spot on the image. This allowed the software to determine the pixel intensity within (foreground) and outside (background) the spot. Before differentially regulated spots can be identified, various other manipulations like normalization and quality control tests have to be carried out.

After the pixel intensities have been determined, the software will proceed to match them with the spot identities and generate a statistical database. Visual analysis can be carried out by comparing the colour saturation and intensities. As shown in Figure 3-18, differentially regulated genes and amount of expression can be picked out by visually analyzing the image. For the case where RNA from virus-infected cells were labelled with Cy5 (red) and RNA from control cells were labelled with Cy3 (green), the spot will be red if the RNA from the infected population is in abundance. If the RNA from the control population is in abundance, the spot will be green. If infected and control RNA samples bind equally, the spot will be yellow. If neither binds, it will not fluoresce and appear black. Thus, from the fluorescence intensities and colours for each spot, the relative expression levels of the genes in the sample and control populations can be estimated. For a closer analysis of an individual spot, the feature viewer can be used to examine the spot features, as shown in Figure 3-19. This will display the wavelength intensity and its ratio for that spot.

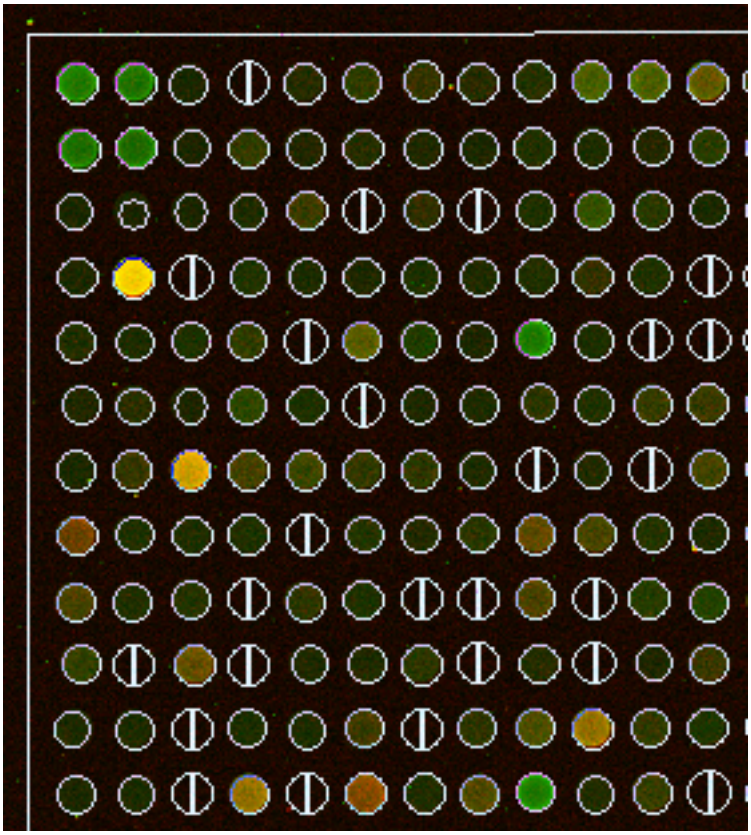


**Figure 3-17.** The microarray image is fitted with a grid of spots, thus allowing the software to determine the positions of the spots and calculate their signal intensities. Spots with a vertical line across denote defective spots as marked out by the slide manufacturer. This is part of their QC process.

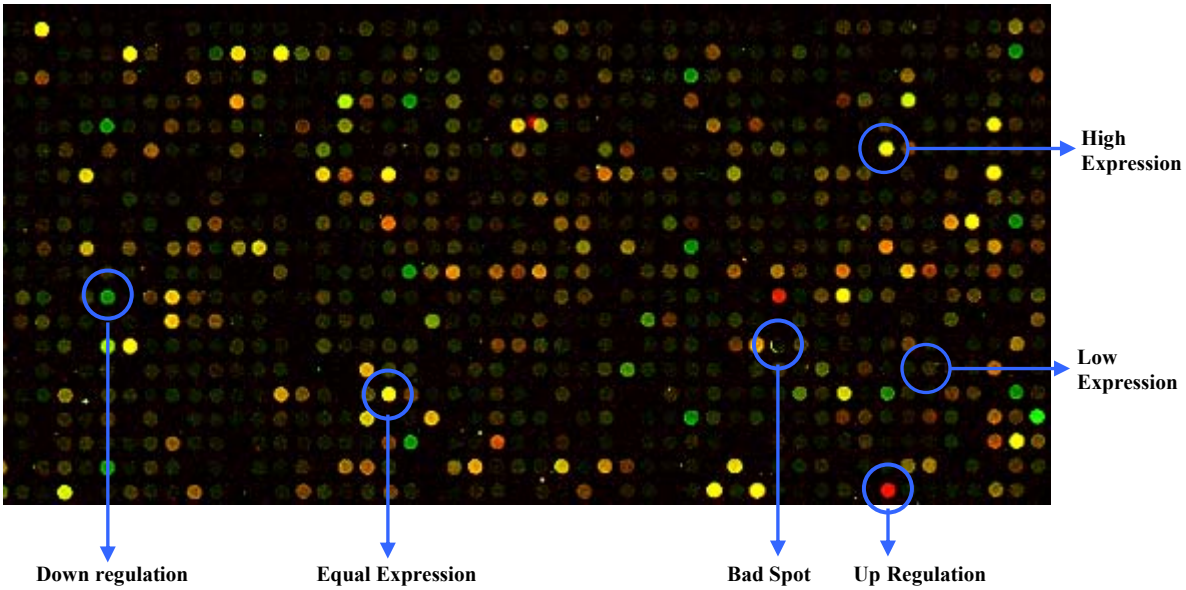



**Figure 3-18.** A sample section of a microarray image, showing differentially regulated genes and bad spots which are omitted from analysis. The diameter of each spot is 135 microns. The above section was taken from HWN2.

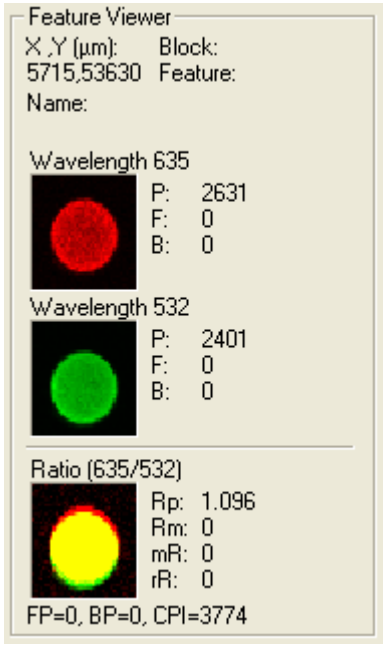
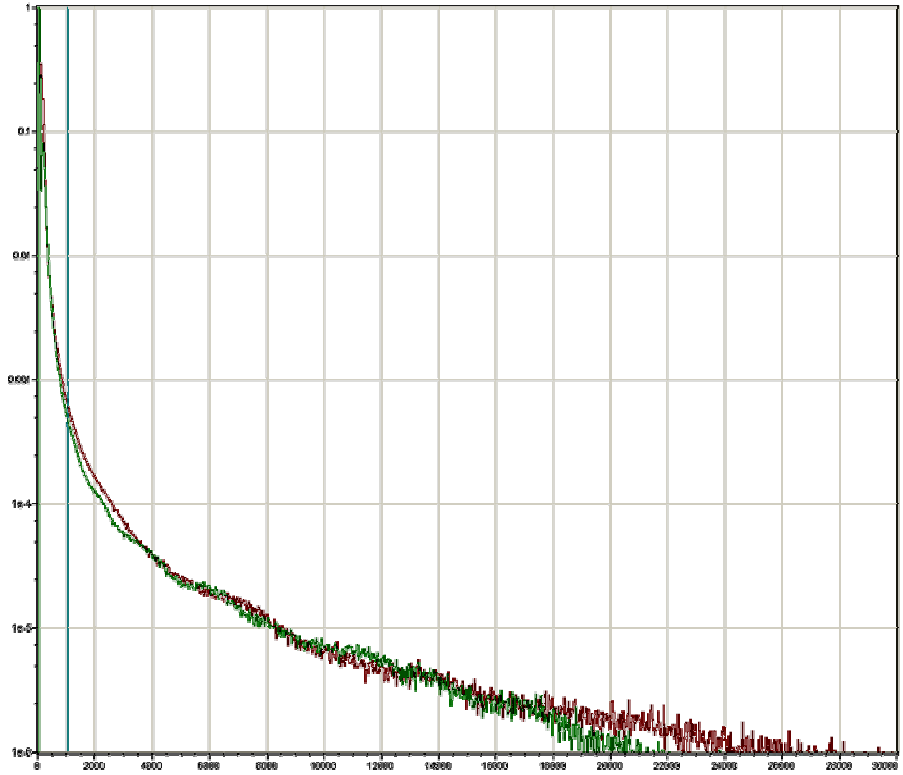



**Figure 3-19.** The feature viewer shows the statistics for a single spot. For example, the Cy5 intensity (top) is 2631 and the Cy3 intensity (middle) is 2401. When the 2 wavelengths are combined, the spot turns up yellow, with a differential expression ratio of 1.096. This gene has approximately equal spot intensities and is therefore equally expressed.

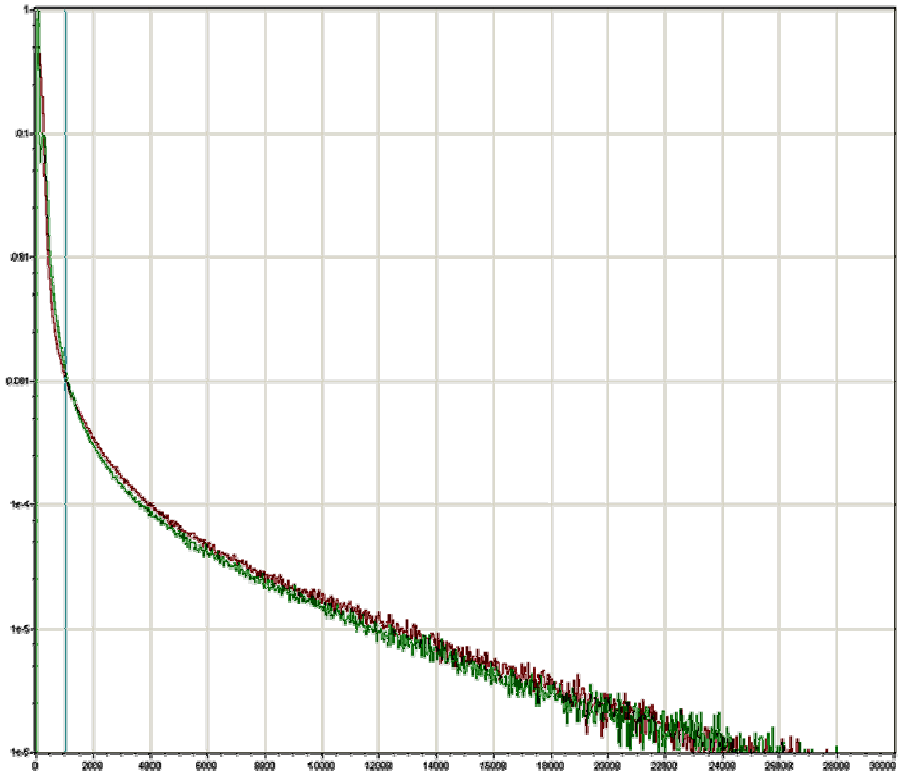
However, before any meaningful information can be gleaned from the mass of data, the intensities for the 2 fluorescent dyes need to be normalized (see Section 2.4.6.3), which corrects for any biases due to the dyes.

A regression method of normalization can be carried out, which tries to equalize the total intensity counts of the two dyes. Figure 3-20 shows the intensity curves before and after normalization. Assuming that the global genomic expression remained constant before and after infection, the total ratio of the Cy3 versus Cy5 should be approximately equal to one. Therefore, after normalization, the green and red curves in Figure 3-20 were approximately equal. A more common method is to utilize Lowess normalization, whereby the values between the 2 dyes are brought together in an intensity dependent manner. Figure 3-21 shows a typical M-A plot of the microarray spots showing the effects after undergoing Lowess normalization. The slight bend of the smoother (trend line) suggested subtle-intensity dependent differential dye bias was present when different dyes were used. A dye-swap duplicate was therefore an essential step here, since incorporation of Cy3 or Cy5 modified nucleotide analogues have been known to manifest a difference in gene-specific incorporation efficiency due to sequence-specific artefacts (Tseng *et al.*, 2001) and dye biases (Holloway *et al.*, 2002).

After normalization, the spots were then ready to be analyzed. Figure 3-22 shows a scatter plot of all the spots. This provided a convenient visual of the differentially regulated spots. There were spots that lie far away from the regression line. Spots that lie on the regression line were equally expressed.



**Before Normalization**



**After Normalization**

**Figure 3-20.** The Intensity Distribution curves for Cy3 (green) and Cy5 (red) are plotted out. The top graph shows the curves before normalization. After normalization, the counts ratio between the 2 dyes are approximately equal to 1, and the green and red curves are approximately equal, as shown by the graph below. The graphs were for sample HWN1.

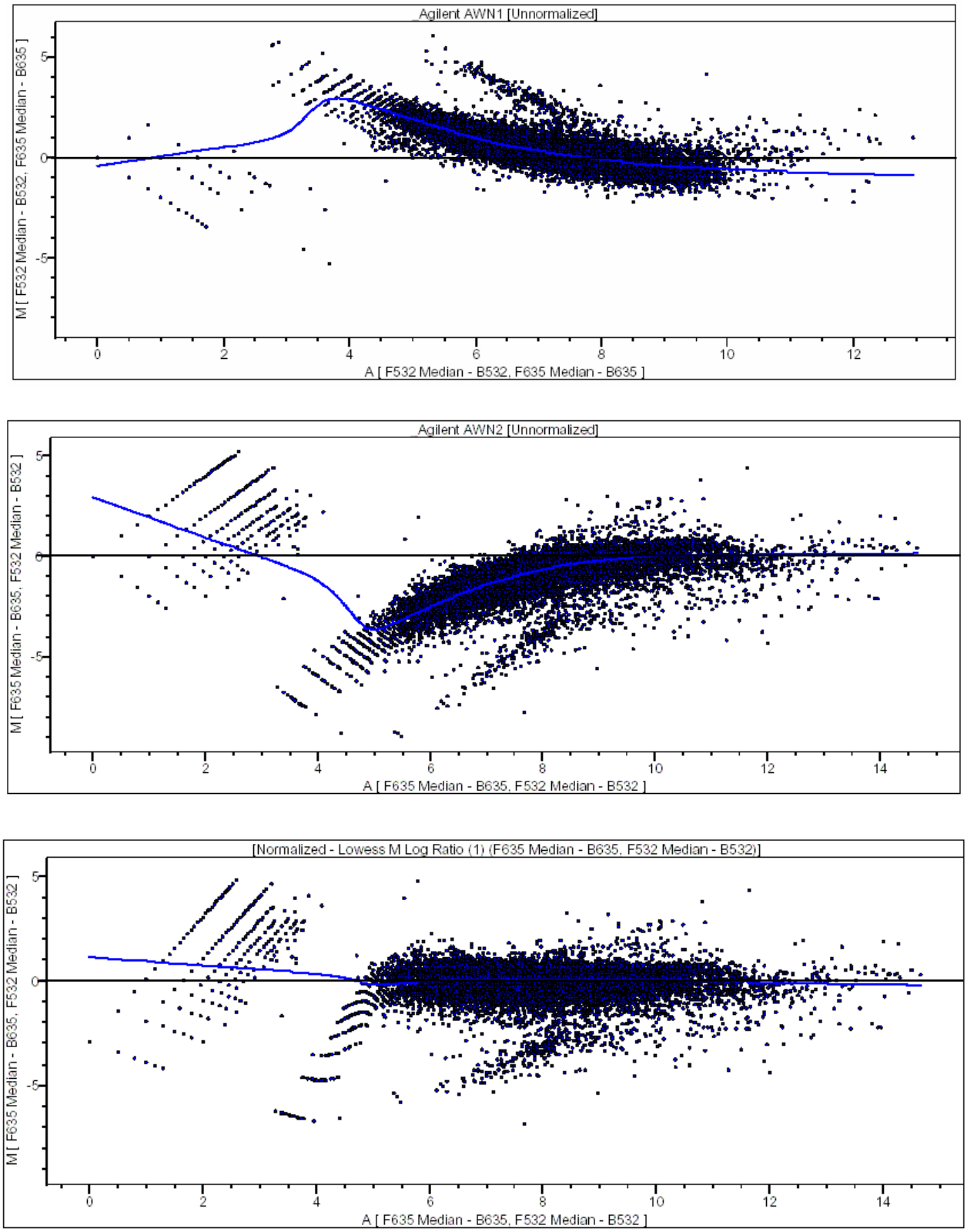



**Figure 3-21.** Intensity-based normalization using the Lowess method. Each spot on the graph corresponds to one feature on the microarray. Spots that are on the zero line are not differentially regulated. The top and middle panel shows the plot from a dye swap experiment. A dye bias is detected from the opposite directions of the deviated smoother curve (trend line). The bottom panel shows the plot after Lowess normalization.

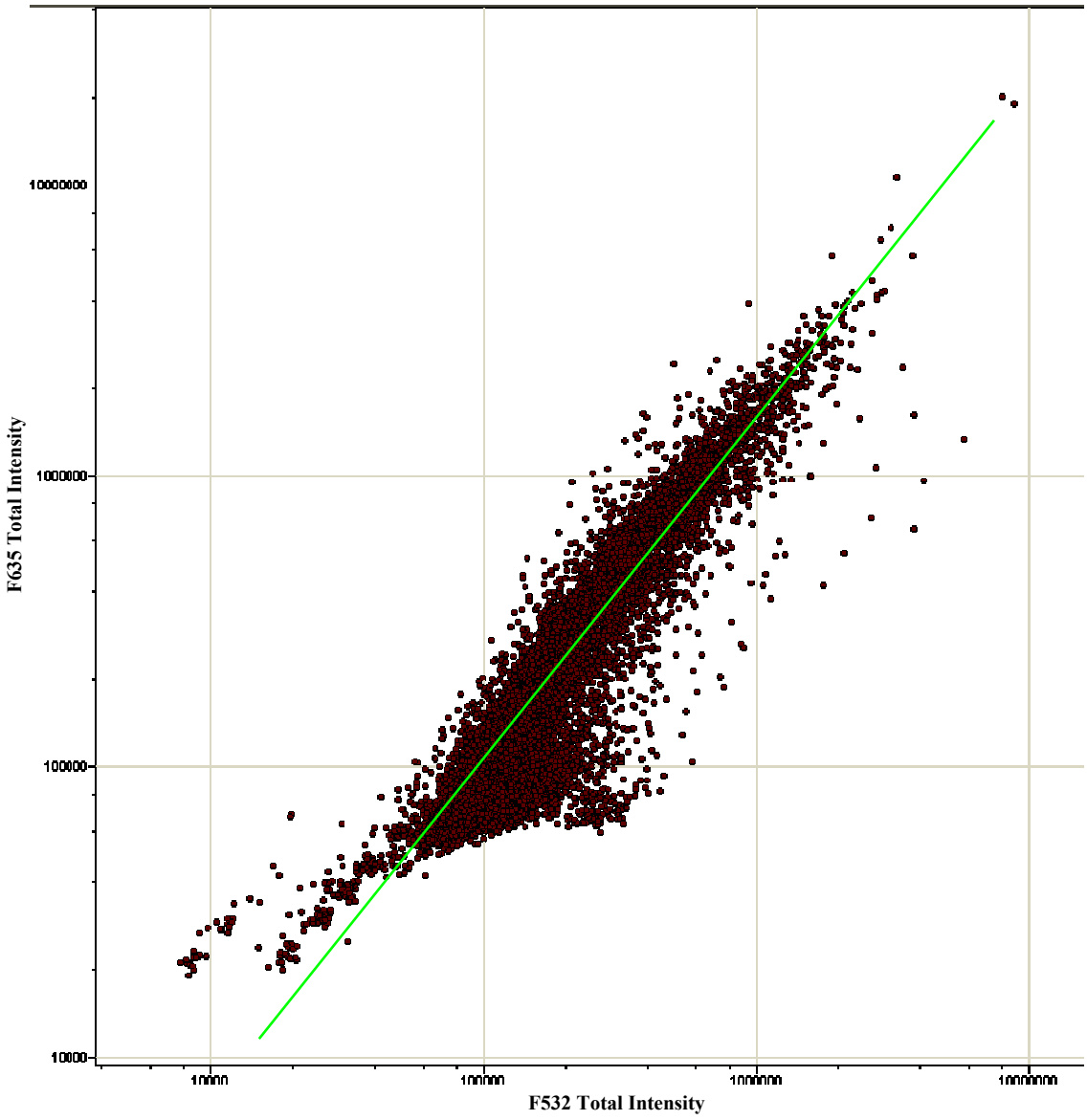



**Figure 3-22.** A scatter plot of the total intensities of every spot on a log graph. The regression line is shown in green. Spots that deviate greatly from this line ($Log_2$ ratios > ±1) are considered to be differentially expressed.

### 3.6.6 Differentially Regulated Genes in West Nile Virus-Infected A172 Cells

A172 cells were found to be more permissive to WNV infection compared to HeLa cells, and this is easy to fathom as A172 cells are of CNS origins, and are thus the natural targets for the WNV. To understand the pathogenesis of brain tissue infected with WNV, an analysis of the global transcriptomics on infected-A172 cells alone was carried out.

A total of 173 cellular genes were identified by ArrayTools to be differentially expressed in the A172 cells after WNV infection, out of which 57 genes were found to be up-regulated and 116 genes were down-regulated. These genes can be found listed in Appendix 6, and they are sorted according to the magnitude of their fold change.

Instead of just sorting these genes into their functional groups, *EASE* (Hosack *et al.*, 2003) was used to identify specific functional groups of genes (or gene ontology) that were found to be highly enriched in occurrence, compared to the whole human genome. The basis for selection is a significance value of less than 0.01 on the *Fisher's* exact *P*-value test. Thirty-nine of the 57 up-regulated genes and 41 of the 116 down-regulated genes were picked out by *EASE*, and these are listed in Tables 3-6 (a and b) according to their *P*-values.

Functional classes that were found to be enriched in the upregulated genes included immune defense, response to external stimulus and pathogens, and apoptosis (Table 3-6a). Genes relating to ubiquitin cycle, transcription regulation and other physiological processes were also identified by *EASE*. Functional classes that were

downregulated were not commonly observed in a virus infection system (Table 3-6b). For instance, genes relating to the mitochondria, ribosomes and protein biosynthesis were found to be highly over-represented in down regulation. Genes which have a putative relevance to influence virus infection are briefly mentioned below.

**Table 3-6a.** List of upregulated functional groups in WN(S)V-infected A172 cells at 24 h post-infection.

| System Category | *Fisher Exact* | *Gene Symb* | *Fold Change* | *GENENAME* |
|---|---|---|---|---|
| **immune defense response** | **1.35E-11** | OAS3 | 2.32 | 2'-5'-oligoadenylate synthetase 3, 100kDa |
| | | GBP5 | 2.47 | guanylate binding protein 5 |
| | | OASL | 3.46 | 2'-5'-oligoadenylate synthetase-like |
| | | HLA-C | 2.20 | major histocompatibility complex, class I, C |
| | | INDO | 3.38 | indoleamine-pyrrole 2,3 dioxygenase |
| | | IFITM1 | 12.00 | interferon induced transmembrane protein 1 (9-27) |
| | | G1P2 | 9.50 | interferon, alpha-inducible protein (clone IFI-15K) |
| | | IFIT2 | 3.76 | interferon-induced protein with tetratricopeptide repeats 2 |
| | | MX2 | 3.17 | myxovirus (influenza virus) resistance 2 (mouse) |
| | | IFIT1 | 10.74 | interferon-induced protein with tetratricopeptide repeats 1 |
| | | IFITM2 | 3.04 | interferon induced transmembrane protein 2 (1-8D) |
| | | IFI27 | 4.03 | interferon, alpha-inducible protein 27 |
| **response to external stimulus** | **7.9E-10** | TRAG3 | 3.53 | taxol resistance associated gene 3 |
| | | SOD2 | 2.12 | superoxide dismutase 2, mitochondrial |
| **ubiquitin cycle** | **0.000871** | CEB1 | 2.33 | cyclin-E binding protein 1 |
| | | FLJ13855 | 2.05 | hypothetical protein FLJ13855 |
| **transcription regulator activity** | **0.0026** | MAZ | 2.02 | MYC-associated zinc finger protein |
| | | TBX3 | 2.32 | T-box 3 (ulnar mammary syndrome) |
| | | KLF2 | 2.62 | Kruppel-like factor 2 (lung) |
| | | ZFP36L2 | 2.16 | zinc finger protein 36, C3H type-like 2 |
| | | NT5C3 | 2.33 | 5'-nucleotidase, cytosolic III |
| | | EGR1 | 4.79 | early growth response 1 |
| **DNA binding** | **0.00528** | KIF22 | 2.02 | kinesin family member 22 |
| | | RBPSUHL | 2.43 | recombining binding protein suppressor of hairless (Drosophila)-like |
| | | SSA1 | 2.29 | Sjogren syndrome antigen A1 |
| **response to pest/pathogen/parasite** | **0.00629** | CXCL10 | 2.10 | chemokine (C-X-C motif) ligand 10 |
| | | CXCL11 | 2.31 | chemokine (C-X-C motif) ligand 11 |
| | | FOSL1 | 2.08 | FOS-like antigen 1 |
| | | PTX3 | 3.44 | pentaxin-related gene, rapidly induced by IL-1 beta |
| **apoptosis** | **0.00947** | TNFSF14 | 2.19 | tumor necrosis factor (ligand) superfamily, member 14 |
| | | BIRC3 | 2.24 | baculoviral IAP repeat-containing 3 |
| | | NFKBIA | 4.13 | nuclear factor of kappa light polypeptide gene enhancer |
| | | TRAF1 | 2.01 | TNF receptor-associated factor 1 |
| | | IFNB1 | 2.79 | interferon, beta 1, fibroblast |
| **physiological process** | **0.00957** | LAP3 | 2.73 | leucine aminopeptidase 3 |
| | | KCNH6 | 2.21 | potassium voltage-gated channel, subfamily H, member 6 |
| | | TFPI2 | 5.34 | tissue factor pathway inhibitor 2 |
| | | CHRND | 2.30 | cholinergic receptor, nicotinic, delta polypeptide |
| **Others** | | RIG-I | 4.19 | DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide |
| | | SAT | 2.18 | spermidine/spermine N1-acetyltransferase |

**Table 3-6b.** List of downregulated functional groups in WN(S)V-infected A172 cells at 24 h post-infection.

| System Category | *Fisher Exact* | *Gene Symb* | *Fold Change* | *GENENAME* |
|---|---|---|---|---|
| **macromolecule biosynthesis** | **5.84E-09** | FDPS | -2.10 | farnesyl diphosphate synthase |
| | | LTA4H | -2.02 | leukotriene A4 hydrolase |
| | | TPI1 | -2.17 | triosephosphate isomerase 1 |
| **Mitochondrion** | **2.22E-07** | ATP5G1 | -2.64 | ATP synthase, mitochondrial F0 complex, subunit c, isoform 1 |
| | | ATP5C1 | -3.82 | ATP synthase, mitochondrial F1 complex, gamma polypeptide 1 |
| | | ATP5J | -2.11 | ATP synthase, mitochondrial F0 complex, subunit F6 |
| | | UQCRB | -2.11 | ubiquinol-cytochrome c reductase binding protein |
| | | COX5B | -2.13 | cytochrome c oxidase subunit Vb |
| | | SLC25A6 | -2.02 | solute carrier family 25 (mitochondrial carrier) |
| | | ATP5B | -2.17 | ATP synthase, mitochondrial F1 complex, beta polypeptide |
| | | ATP5A1 | -2.21 | ATP synthase, mitochondrial F1 complex, alpha subunit, isoform 1 |
| | | SDHC | -2.31 | succinate dehydrogenase complex, subunit C |
| | | ATP5O | -2.00 | ATP synthase, mitochondrial F1 complex, O subunit |
| | | PRDX5 | -2.74 | Peroxiredoxin 5 |
| | | BNIP3L | -2.49 | BCL2/adenovirus E1B 19kDa interacting protein 3-like |
| | | COX6B | -2.41 | cytochrome c oxidase subunit VIb |
| | | ATP5F1 | -2.43 | ATP synthase, mitochondrial F0 complex, subunit b, isoform 1 |
| | | COX7B | -2.21 | cytochrome c oxidase subunit VIIb |
| **cytosolic ribosome (sensu Eukarya)** | **4.97E-07** | RPS3A | -3.33 | ribosomal protein S3A |
| | | RPLP0 | -2.16 | ribosomal protein, large, P0 |
| | | RPL4 | -2.30 | ribosomal protein L4 |
| | | RPL41 | -2.03 | ribosomal protein L41 |
| | | RPS24 | -2.51 | ribosomal protein S24 |
| | | RPL7A | -2.03 | ribosomal protein L7a |
| | | RPS3A | -3.04 | ribosomal protein S3A |
| | | RPL5 | -2.97 | ribosomal protein L5 |
| | | RPS3A | -2.23 | ribosomal protein L23 |
| | | RPL26L1 | -2.12 | ribosomal protein L26-like 1 |
| **protein bio-synthesis** | **1.63E-06** | NACA | -2.17 | nascent-polypeptide-associated complex alpha polypeptide |
| | | EIF3S5 | -2.88 | eukaryotic translation initiation factor 3, subunit 5 epsilon, 47kDa |
| | | SRP9 | -3.27 | signal recognition particle 9kDa |
| | | EIF4G2 | -2.11 | eukaryotic translation initiation factor 4 gamma, 2 |
| | | EEF2 | -2.60 | eukaryotic translation elongation factor 2 |
| **cytoskeletal (actin) binding** | **0.00111** | CAPZA2 | -2.12 | capping protein (actin filament) muscle Z-line, alpha 2 |
| | | CNN3 | -2.50 | calponin 3, acidic |
| | | MYH6 | -3.00 | myosin, heavy polypeptide 6, cardiac muscle, alpha |
| | | DSTN | -2.13 | destrin (actin depolymerizing factor) |
| **small GTPase mediated signal transduction** | **0.00367** | RAC1 | -2.87 | ras-related C3 botulinum toxin substrate 1 |
| | | ARF4 | -2.30 | ADP-ribosylation factor 4 |
| | | ARFD1 | -2.20 | ADP-ribosylation factor domain protein 1, 64kDa |
| | | ARHI | -2.73 | ras homolog gene family, member I |
| **Others** | | EEF1G | -3.29 | eukaryotic translation elongation factor 1 gamma |
| | | S100A4 | -2.53 | S100 calcium binding protein A4 |
| | | HDAC3 | -2.52 | histone deacetylase 3 |
| | | PFN2 | -2.42 | profilin 2 |
| | | PRDX3 | -2.27 | Peroxiredoxin 3 |
| | | KRT15 | -2.18 | keratin 15 |
| | | LASP1 | -2.14 | LIM and SH3 protein 1 |

Of the genes found in the immune defense class, OAS3 and MX2 have known antiviral properties. OAS3 catalyzes the 2', 5' oligomers of adenosine in order to bind and activate RNase L, while MX2 is a member of both the dynamin family and the family of large GTPases. Genes like CXCL10 and CXCL11 belong to the CXC subfamily of chemokines that modulates the immune response. Other immune related genes that showed an increase in expression include HLA-C, which codes for the MHC-I, and PTX3, which is an acute phase protein. Both INDO and FOSL1 genes have an effect on the regulation of cell proliferation and transformation. A number of interferon-induced proteins (e.g. IFIT1, IFI27, IFITM1) were found to be up-regulated, but their functions are not known, and may suggest novel roles in inhibiting virus replication. SOD2 gene codes for superoxide dismutase, which is involved in protection from free radicals.

Many of the genes in the transcription regulation group relates to growth regulation. Examples include EGR1 and ZFP36L2 genes. CEB1 gene belongs to the ubiquitin cycle functional group and also helps to regulate growth. Genes associated with apoptosis were also found to have increased in expression. These include TNFSF14, TRAF1 and NFKBIA genes, which are proapoptotic, and BIRC3 gene, which codes for a baculoviral inhibitor of apoptosis proteins (IAP). SAT is an enzyme that catalyzes the N(1)-acetylation of spermidine and spermine, and may also be involved in apoptosis (Babbar *et al.*, 2003).

Genes whose proteins are able to bind DNA include KIF22, which is a kinesin that has been found to assist in the transport of organelles by also binding to microtubules (Shiroguchi *et al.*, 2003) The RBPSUHL gene is a transcription factor to activate

transcription in concert with Epstein-Barr virus nuclear antigen-2 (EBNA2) (Minoguchi *et al.*, 1997).

Other genes of interest included TFPI2, which is a serine protease inhibitor belonging to the physiological processes functional category. RIG-I belongs to the DEAD box family of proteins and encodes for RNA helicases.

Genes associated with cell metabolism was found to be greatly decreased. FDPS is involved in the isoprene biosynthetic pathway that provides the cell with cholesterol, ubiquinone, dolichol, and other nonsterol metabolites. TPI1 is a triosephosphate isomerase and its deficiency has been implicated in neurodegenerative diseases (Olah *et al.*, 2002).

Many genes associated with the mitochondria were also found to be reduced in expression. SDHC, COX5B, 6B, and 7B are genes involved in the mitochondrial respiratory chain. Some mitochondrial genes involved in cellular protection were also found to be down-regulated. These were PRDX5 gene, which is an antioxidant enzyme, and BNIP3L gene, which plays a role in apoptosis and tumour suppression. PRDX3 gene, which paradoxically was not classified together in this functional class, was found to be downregulated and is required for normal mitochondria function (Wonsey *et al.*, 2002). A host of genes relating to ATP synthase, which is involved in proton transport, were also found to be down-regulated.

The ribosomal proteins represent the next significant functional category that was down-regulated, consisting of 10 ribosomal proteins belonging to both large and

small subunits. RPL5 protein binds 5S rRNA to form a stable complex called the 5S ribonucleoprotein particle (RNP), which is necessary for the transport of nonribosome-associated cytoplasmic 5S rRNA to the nucleolus for assembly into ribosomes. One of the genes, RPS3A, belongs to the S3AE family of ribosomal proteins and down regulation of this gene was previously found to be involved in cellular differentiation (Goodin & Rutherford, 2002).

Protein biosynthesis represents the next functional class that was found to be down regulated. Translation initiation and elongation factors (EIF3S5, EIF4G2, and EEF2) were found to be represented in this class. The EEF2 gene is an essential factor for protein synthesis by promoting the GTP-dependent translocation of the nascent protein chain from the A-site to the P-site of the ribosome. The EEF1G gene encodes a subunit of the elongation factor-1 complex, which is responsible for the enzymatic delivery of aminoacyl tRNAs to the ribosome. NACA protein, which is part of a heterodimeric complex of alpha- and beta-subunits that prevents mistargeting of nascent polypeptide chains to the endoplasmic reticulum membranes, also belongs to this class.

Actin binding proteins were also found to be highly represented. Genes in this class include DSTN which helps in actin depolymerization, CAPZA2 which is a capping protein at the barbed end of an actin filament, and CNN3 which plays a role in the cellular organization of actin filaments. Genes represented in the small GTPase mediated signal transduction group appear to regulate a diverse array of cellular activities. Of interest is the RAC1 gene which regulates cytoskeletal reorganization

(Shin *et al.*, 2004), and both ARF4 and ARFD1 which plays a role in the formation of intracellular transport vesicles and vesicular trafficking.

Genes that were not classified into functional categories, but were found to be of interest in this study, are also listed. The HDAC3 gene codes for histone deacetylase 3 that plays a critical role in transcriptional regulation, cell cycle progression, and developmental events. Down expression of HDAC3 has been implicated in increasing cell permissiveness to human cytomegalovirus (Murphy *et al.*, 2002). The S100A4 gene is a member of the S100 family of proteins containing 2 EF-hand calcium-binding motifs, and involved in the regulation of a number of cellular processes such as cell cycle progression and differentiation. It has also been recently found to affect F-actin polymerization (Belot *et al.*, 2002), in association with PFN2 gene, which is also down regulated. Other genes that were associated with the cytoskeleton, which were also similarly down regulated, include KRT15 which is a cytokeratin involved in intermediate filament formation, and LASP1 which is involved in cytoskeletal organization and cell motility (Butt *et al.*, 2003).

### 3.6.7 Differentially Regulated Genes between West Nile Virus-Infected A172 and HeLa Cells

A comparative study between A172 and HeLa cells was performed to determine the molecular mechanisms that might explain the difference in permissiveness of the two different cell lines to WN(S) virus infection. Differences in host gene expression patterns between the two cells lines during infection are very likely to have an effect on the efficiency of virus replication. On the other hand, similarities in host genes expression patterns between the two cell lines will have the same effects on the virus

ability to replicate within the cells, and are therefore not listed. For example, signs of apoptosis can be detected in both A172 and HeLa cells, and genes related to apoptosis were upregulated in both cell lines. Apoptosis genes therefore do not show differences in expression patterns between the two cell lines, and are therefore not detected by the software during analysis.

ArrayTools identified 300 cellular genes that showed a 2-fold difference in expression between the two cell lines. For example, if a gene shows a +½ fold increase in A172 cells but shows a –½ fold decrease in HeLa cells, this gene will be identified as showing a 2-fold difference between the two cells lines. After clustering of these genes into their specific functional groups and uploaded into *EASE*, 46 of these genes belonging to 4 functional classes were found to be over-represented, based on a significance value of less than 0.01 on the *Fisher's* exact *P*-value test. These genes are listed in Table 3-7. The individual changes in gene expression within the cell lines are shown, before the fold change between both cells lines are calculated.

Genes relating to intracellular structure and transport were highly represented. Genes from the hexose metabolism group were also present. This was surprising as these genes are highly invariable and are known as housekeeping genes. As in Table 3-6b on A172 cells only, genes relating to protein metabolism and RNA processing were again found to be significantly differentially regulated. The possible relevance of these genes on virus permissiveness can be found in the discussion.

**Table 3-7.** List of differentially expressed genes between WN(S)V-infected A172 and HeLa cells.

| *Classification* | *Symbol* | *Gene Name* | *A172* | *HeLa* | *Fold* |
|---|---|---|---|---|---|
| Intracellular Structure and Transport | DNAH3 | dynein, axonemal, heavy polypeptide 3 | 1.06 | -0.53 | 3.02 |
| | TUBA2 | tubulin, alpha 2 | 0.85 | -0.62 | 2.76 |
| | STX4A | syntaxin 4A (placental) | 0.68 | -0.44 | 2.17 |
| | C15orf22 | integral type I protein | 0.40 | -0.64 | 2.06 |
| | ARFGAP3 | ADP-ribosylation factor GTPase activating protein 3 | 0.23 | -0.91 | 2.21 |
| | IPO8 | importin 8 | 0.11 | -1.55 | 3.14 |
| | SLC25A6 | solute carrier family 25, member 6 | -0.01 | -1.17 | 2.24 |
| | ANK3 | ankyrin 3 | -0.02 | -1.11 | 2.12 |
| | PSAP | prosaposin | 0.16 | -1.74 | 3.72 |
| | FRMD4A | FERM domain containing 4 | 0.11 | -3.11 | 9.31 |
| | KLC2L | kinesin light chain 2-like | 1.34 | -0.84 | 4.53 |
| | CKAP4 | cytoskeleton-associated protein 4 | 0.81 | -0.46 | 2.40 |
| | KRTAP2 | keratin associated protein 2-1 | 0.57 | -1.03 | 3.03 |
| | KRTHA3B | keratin, 3B | 0.36 | -0.76 | 2.18 |
| | K-ALPHA-1 | tubulin, alpha | 0.36 | -0.93 | 2.45 |
| | ACTN4 | actinin, alpha 4 | 0.23 | -1.00 | 2.34 |
| | NXF5 | nuclear RNA export factor 5 | -0.10 | 1.32 | 0.37 |
| | TEKT3 | tektin 3 | 0.18 | 2.37 | 0.22 |
| | CENPJ | centromere protein J | -0.16 | 1.01 | 0.45 |
| | PFN2 | profilin 2 | -0.84 | 0.50 | 0.40 |
| | TUBB2 | tubulin, beta, 2 | -0.99 | 0.36 | 0.39 |
| | ACTG1 | actin, gamma 1 | -0.33 | 0.89 | 0.43 |
| Hexose Meta-bolism | ENO1 | enolase 1, (alpha) | 0.53 | -1.14 | 3.20 |
| | ALDOA | aldolase A, fructose-bisphosphate | 0.42 | -0.74 | 2.24 |
| | G6PD | glucose-6-phosphate dehydrogenase | 0.07 | -1.39 | 2.76 |
| | SUCLG2 | succinate-CoA ligase, GDP-forming | 0.00 | -1.01 | 2.01 |
| Protein Metabolism | SRM | spermidine synthase | 0.32 | -0.80 | 2.18 |
| | INDO | indoleamine-pyrrole 2,3 dioxygenase | 2.98 | -0.07 | 8.28 |
| | WARS | tryptophanyl-tRNA synthetase | 1.75 | 0.44 | 2.47 |
| | EEF1A2 | eukaryotic translation elongation factor 1 alpha 2 | -0.06 | -1.09 | 2.04 |
| | EEF1G | eukaryotic translation elongation factor 1 gamma | -0.53 | -1.66 | 2.18 |
| | TUFM | Tu translation elongation factor, mitochondrial | 0.56 | -0.76 | 2.50 |
| | RPLP2 | ribosomal protein, large P2 | -0.19 | -1.82 | 3.09 |
| | RPL10L | ribosomal protein L10-like | -0.51 | -1.79 | 2.43 |
| | RPS24 | ribosomal protein S24 | -0.66 | 0.42 | 0.47 |
| | RPS13 | ribosomal protein S13 | -0.60 | 0.58 | 0.44 |
| RNA Processing | SNRPB | small nuclear ribonucleoprotein polypeptides B and B1 | 0.65 | -1.08 | 3.33 |
| | FNBP3 | formin binding protein 3 | 0.60 | -0.54 | 2.19 |
| | TRA2A | transformer-2 alpha | 0.40 | -0.67 | 2.10 |
| | SF3B5 | splicing factor 3b, subunit 5, 10kDa | 0.36 | -0.96 | 2.49 |
| | SURF6 | surfeit 6 (localize to nucleolus) | 0.19 | -0.98 | 2.25 |
| | RNASE6 | ribonuclease, RNase A family, k6 | -0.37 | 0.75 | 0.46 |
| | SF4 | splicing factor 4 | -0.93 | 0.27 | 0.44 |
| | NXF5 | nuclear RNA export factor 5 | -0.10 | 1.32 | 0.37 |

**A172, HeLa** = gene expression fold change of infected cells vs. mock-infected cells, expressed in Log2.
**Fold** = gene expression fold change of infected-A172 cells vs. infected-HeLa cells.

### 3.6.8 Confirmation of Expression Changes by Quantitative Real-Time PCR (qRT-PCR) Analysis.

The qRT-PCR was performed on a fresh set of RNA from that used for the microarray experiments to confirm the microarray results. This was to ensure that the genes identified were truly differentially expressed due to the virus infection. Genes for qPCR were chosen such that a broad spectrum of functional classes was represented. From Table 3-8, confirmation tests on qPCR corroborated with microarray results, thereby verifying the accuracy of the statistical analyses. Two genes (DUSP1 and DNAJB1) that showed less than 2-fold change on the microarray, were found to be differentially expressed to a greater extent, thus substantiating the sensitivity of the microarray tests. Fold changes from qRT-PCR were observed to be mostly of a higher magnitude compared to microarray results, and this may be due to the high background or saturation of the fluorescence signal in the microarrays.

**Table 3-8.** Comparison of gene expression changes between microarray and qRT-PCR

| Gene Syb | Gene Name | Microarray fold change | RT-PCR fold change |
|---|---|---|---|
| ARHI | ras homolog gene family | -2.72 | -2.55 |
| ATP5J | ATP synthase, mitochondrial F0 complex | -2.11 | -2.60 |
| CEB1 | hect domain and RLD 5 | 2.32 | 42.22 |
| DNAJB1 | DnaJ (Hsp40) homolog, subfamily B | -1.97 | -2.14 |
| DUSP1 | dual specificity phosphatase 1 | 1.92 | 5.66 |
| EGR1 | early growth response 1 | 4.79 | 8.57 |
| EIF4G2 | eukaryotic translation initiation factor 4 gamma, 2 | -2.11 | -7.77 |
| FLJ13855 | hypothetical protein FLJ13855 | 2.05 | 3.85 |
| FOSL1 | FOS-like antigen 1 | 2.08 | 6.50 |
| IFITM1 | interferon induced transmembrane protein 1 (9-27) | 12.03 | 527.61 |
| LTA4H | leukotriene A4 hydrolase | -2.02 | -8.10 |
| RPL5 | ribosomal protein L5 | -2.97 | -9.03 |
| RPL7A | ribosomal protein L7a | -2.03 | -3.42 |
| RPLP0 | ribosomal protein, large, P0 | -2.15 | -1.52 |
| TFPI2 | tissue factor pathway inhibitor 2 | 5.21 | 11.58 |

# 4.0 RESULTS – Progressive Host Interactions with West Nile Virus during Infection

In Chapter 3, results of the microarray studies between A172 and HeLa cells were presented. A 'snap-shot' of the host gene regulation of the infected cells was obtained for 24 h p.i., which corresponds to the late phase of infection or peak virus production. As A172 cells were found to be more suitable for WN(S)V infection, it was used for further studies. In this chapter, the results of a time-course study on the changes in global transcriptomics on A172 cells are presented. The aim is to determine the aberrations in gene expression as the virus infection progressed. The time points used were 1½ h, 6 h, 12 h, 18 h, and 24 h p.i. The choice of these time points allows us to understand the changes in host gene response starting from the early to the late phases of infection.

## 4.1 Preparation of Samples for Microarray Studies

As the quality of scanned images resulting from the use of Amersham's kits and hybridization stations tend to yield inconsistent results, it was decided to switch to Agilent Technologies' recommended kits that are more appropriate to the microarray slides used here. As it was difficult to assess if the inconsistency was with the probe labelling or hybridization step, both of these methods were changed and a detailed description of these can be found in Section 2.4.5.

Table 4-1 gives the values of the quantity of total RNA in the samples used in this procedure. The samples were obtained from A172 cells infected after 1.5 h, 6 h, 12 h,

and 18 h p.i. Mock-infected control cells were prepared concurrently at the same timings. Microarray results from the 24 h p.i. time-point were used from the previous section for the microarray analysis here. As the use of the Linear Amplification kit generates cRNA (instead of cDNA), the amount of product generated was measured in terms of the concentration of cRNA (instead of quantity of CyDye). The product quantities are listed in Table 4-2. Figure 4-1 shows a microarray image obtained by using the protocol from Agilent Technologies. As can be seen, images generated by using this protocol are generally of a much higher quality. It shows a wider dynamic range in spot intensities, with clear background and absence of any blotchy patterns.

**Table 4-1:** Quantity and Purity of RNA samples.

| Sample | | Dye | $A_{260}$ | $A_{280}$ | Concentration (µg/µl) | Yield (µg) | Purity |
|---|---|---|---|---|---|---|---|
| **A1.5WN1** | **A1.5WNV1** | Cy3 | 0.475 | 0.216 | 0.950 | 47.5 | 2.20 |
| | **A1.5WNC1** | Cy5 | 0.624 | 0.293 | 1.248 | 62.4 | 2.13 |
| **A1.5WN2** | **A1.5WNV2** | Cy5 | 0.816 | 0.373 | 1.632 | 81.6 | 2.19 |
| | **A1.5WNC2** | Cy3 | 0.738 | 0.369 | 1.476 | 73.8 | 2.00 |
| **A6WN1** | **A6WNV1** | Cy3 | 0.965 | 0.448 | 1.930 | 96.5 | 2.15 |
| | **A6WNC1** | Cy5 | 0.853 | 0.436 | 1.706 | 85.3 | 1.96 |
| **A6WN2** | **A6WNV2** | Cy5 | 0.824 | 0.461 | 1.648 | 82.4 | 1.79 |
| | **A6WNC2** | Cy3 | 0.820 | 0.420 | 1.640 | 82.0 | 1.95 |
| **A12WN1** | **A12WNV1** | Cy3 | 1.055 | 0.506 | 2.110 | 105.5 | 2.08 |
| | **A12WNC1** | Cy5 | 1.329 | 0.653 | 2.658 | 132.9 | 2.04 |
| **A12WN2** | **A12WNV2** | Cy5 | 0.954 | 0.441 | 1.908 | 95.4 | 2.16 |
| | **A12WNC2** | Cy3 | 0.853 | 0.410 | 1.706 | 85.3 | 2.08 |
| **A18WN1** | **A18WNV1** | Cy3 | 0.734 | 0.361 | 1.468 | 73.4 | 2.03 |
| | **A18WNC1** | Cy5 | 1.047 | 0.478 | 2.094 | 104.7 | 2.19 |
| **A18WN2** | **A18WNV2** | Cy5 | 0.854 | 0.411 | 1.708 | 85.4 | 2.08 |
| | **A18WNC2** | Cy3 | 0.805 | 0.394 | 1.610 | 80.5 | 2.04 |

**Sample:** A – A172, WNV – Virus infected sample, WNC – Control sample, (1.5, 6, 12, 18) – represents the number of h p.i. (24 h p.i. was not carried out as the raw microarray results were obtained from the previous section.)

**$A_{260}$/$A_{280}$:** UV Absorbance readings.

**Concentration** = $A_{260}$ X 40 X dilution factor (50)

**Yield** = Concentration X volume (50)

**Purity** = $A_{260}/A_{280}$

**Table 4-2:** Quantity of cRNA generated.

| Sample | | $A_{260}$ | [cRNA] (µg/µl) |
|---|---|---|---|
| **A1.5WN1** | **A1.5WNV1** | 0.056 | 44.8 |
| | **A1.5WNC1** | 0.058 | 46.4 |
| **A1.5WN2** | **A1.5WNV2** | 0.075 | 60.0 |
| | **A1.5WNC2** | 0.089 | 71.2 |
| **A6WN1** | **A6WNV1** | 0.056 | 44.8 |
| | **A6WNC1** | 0.055 | 44.0 |
| **A6WN2** | **A6WNV2** | 0.067 | 53.6 |
| | **A6WNC2** | 0.066 | 52.8 |
| **A12WN1** | **A12WNV1** | 0.033 | 26.4 |
| | **A12WNC1** | 0.026 | 20.8 |
| **A12WN2** | **A12WNV2** | 0.024 | 19.2 |
| | **A12WNC2** | 0.021 | 16.8 |
| **A18WN1** | **A18WNV1** | 0.026 | 20.8 |
| | **A18WNC1** | 0.026 | 20.8 |
| **A18WN2** | **A18WNV2** | 0.019 | 15.2 |
| | **A18WNC2** | 0.026 | 20.8 |

**Sample:** A – A172, WNV – Virus infected sample, WNC – Control sample, (1.5, 6, 12, 18) – represents the number of h p.i. (24 h p.i. was not carried out as the raw microarray results were obtained from the previous section.)

**$A_{260}$:** UV Absorbance readings.

**Concentration** = $A_{260}$ X 40 X dilution factor (20)

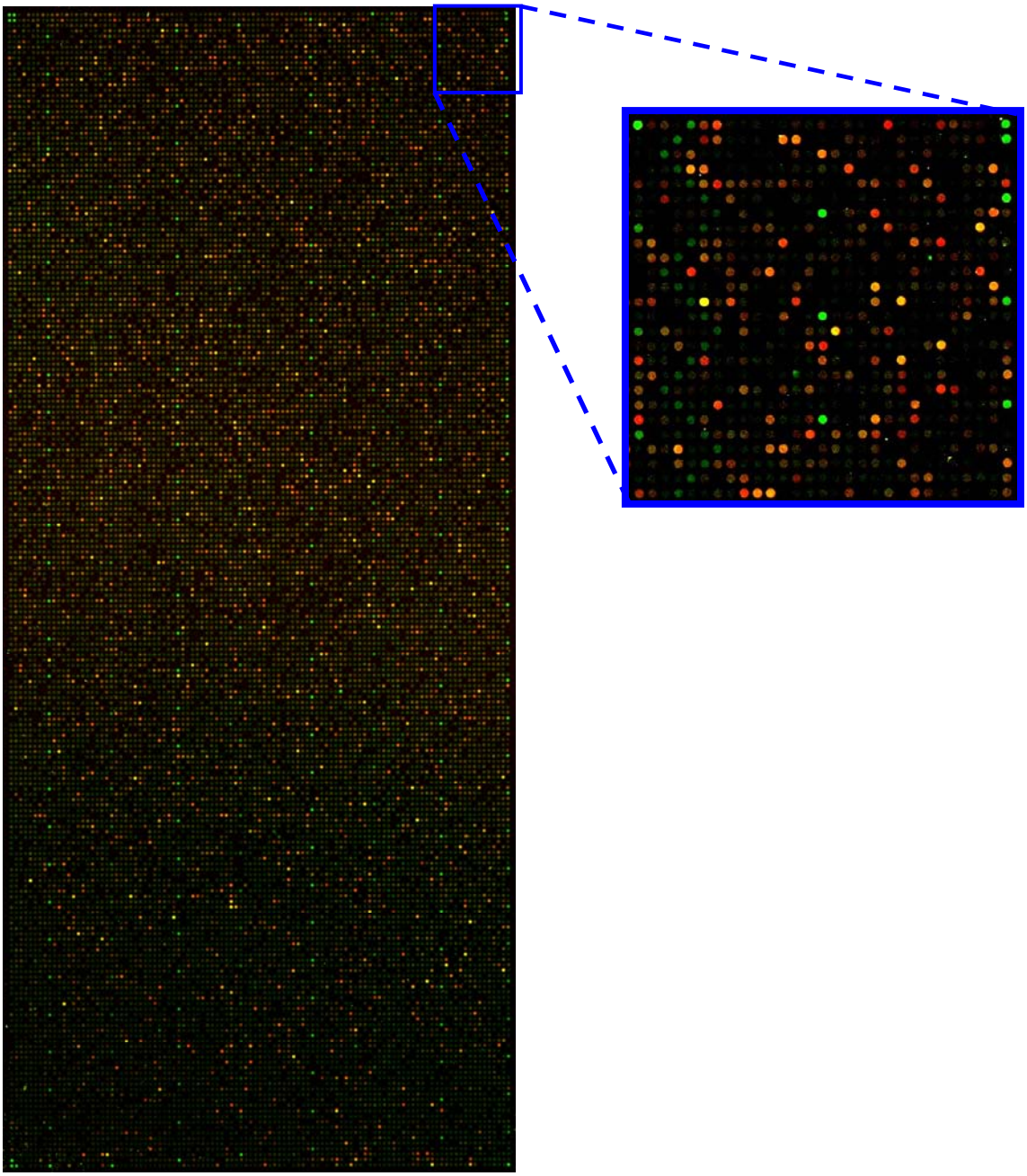


**Figure 4-1.** A scanned microarray image obtained by using the protocol from Agilent Technologies. The blue square shows the magnified corner of the microarray.

## 4.2 Data Transformation from the Raw Data

After the microarrays were scanned and the spot intensities measured for all the slides, the raw data files were fed into the MIDAS software (Saeed *et al.*, 2003) from the Institute of Genomic Research (TIGR, USA). This software carries out the preliminary data processing to normalize the spot intensities, compare the consistency of the spot intensities between replicates and flip-dye duplicates, as well as to filter out differentially regulated genes from each microarray slide. As the preliminary data transformation carried out was the same for all the slides, data from the A12WN slides will be presented here as the representative sample. The data generated from the other microarray slides show similar trends.

As before, the first step in data transformation involved the normalization of the spot intensities using the Lowess method. Figures 4-2 and 4-3 show the effects of the normalization on the spot intensities using this method. After normalization, all the spots were centred on the zero line. The further away the spot was from the zero line, the greater its regulation.

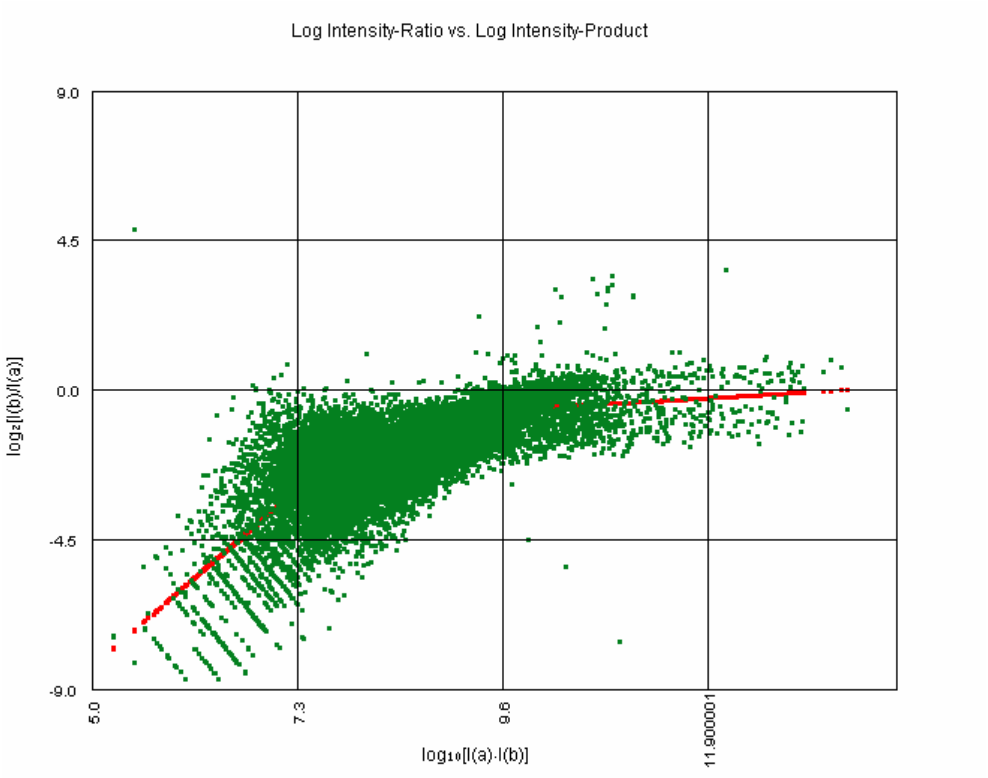



**Figure 4-2**. Pre-Lowess normalization for A12WN1.

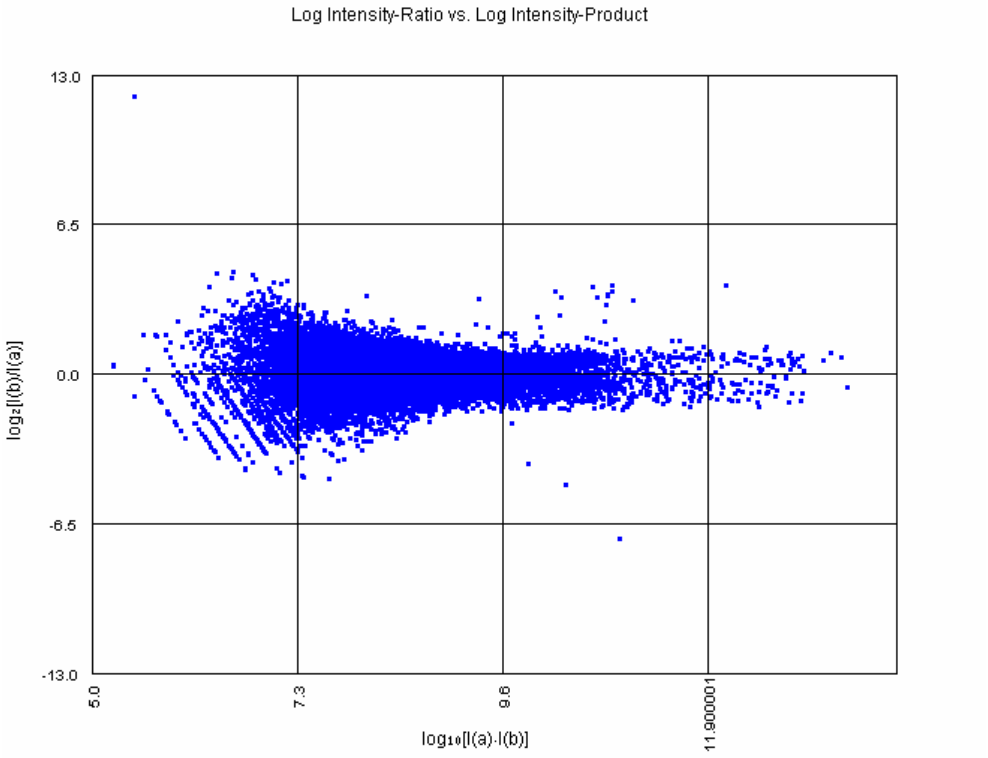



**Figure 4-3.** Post-Lowess normalization for A12WN2.

After normalization, the data between the flip-dye experiments were compared for consistency. The ratio of the spot intensity between the two different dyes should be about the same when flipped in the two microarray slides, and the spots should fall close to the diagonal line when plotted on graph plot (shown in red spots). However, if the spot shows a difference of greater than 2 standard deviations from the mean, then the spots will be filtered out (shown in blue spots). Figure 4-4 shows the typical graph generated from this filtering process. Spots shown in red falls within the 2 standard deviations, and were therefore carried over to the next stage of data analyses. Table 4-3 shows the results of all the slides from this process. During the flip-dye consistency checking process, the values from both slides were combined and a single data file was generated per pair of duplicate experiments.

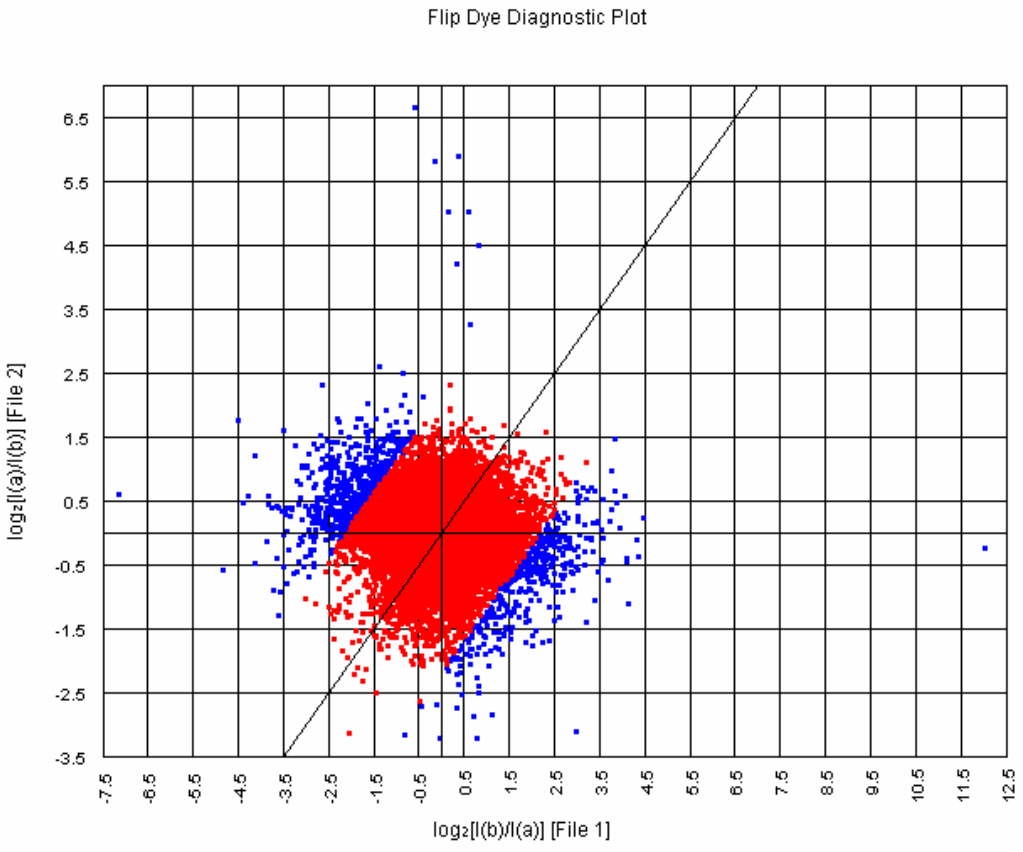



**Figure 4-4.** Flip-dye consistency checking of spots.

Following the flip-dye consistency checking, the data was filtered for differentially regulated genes. This was performed using the Slice Analysis function within MIDAS. In this case, differentially regulated genes are based on their z-score, which is defined as 1.50 standard deviations away from the mean of the intensities. These genes are marked in red in Figure 4-5. As can be seen in the graph plot, spots with low intensities show a wider spread compared to spots with high intensities. This wider spread is attributed to the higher error rates due to ratio comparisons of low values. The *z*-score method therefore allows the identification of differentially regulated genes of more than 2-fold change in a more accurate manner. The number of genes that passed this test is also listed in Table 4-3.

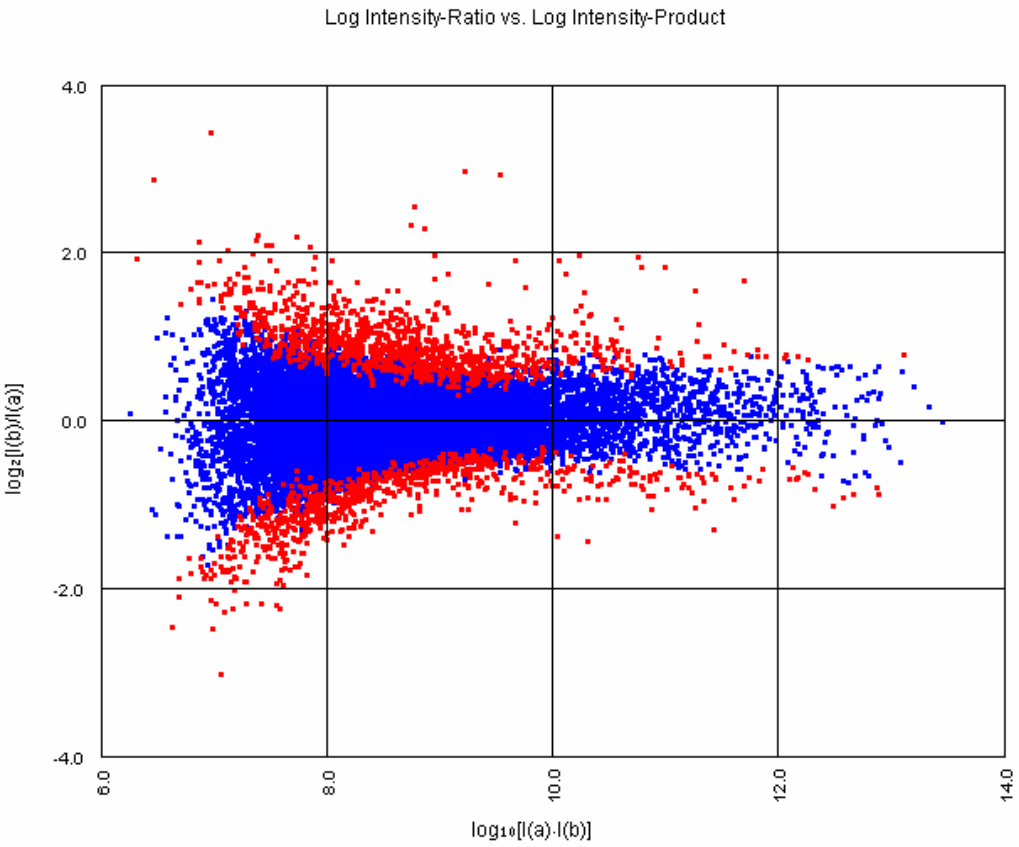



**Figure 4-5.** *z*-score slice analysis showing differentially regulated genes (in red) for the A12WN data.

**Table 4-3.** Results obtained from flip-dye consistency checking and *z*-score slice analysis.

| Sample | Pre-Flip-dye Check | Post-Flip-dye Check | % Passed Filter | Dispersion Factor | Differentially regulated genes |
|---|---|---|---|---|---|
| **A1.5WN** | 16691 | 15762 | 5.57% | 0.335 | 1904 |
| **A6WN** | 17050 | 16194 | 5.02% | 0.238 | 1955 |
| **A12WN** | 16005 | 15008 | 6.23% | 0.289 | 1855 |
| **A18WN** | 16881 | 15220 | 9.84% | 0.604 | 1394 |
| **A24WN** | 16043 | 15211 | 5.19% | 0.303 | 1785 |

## 4.3 Analysis of the Microarray Data

After the differentially regulated genes for the five time points had been identified, they were loaded into TIGR's MultiExperiment Viewer (TMEV) for more advanced data analyses. Hierarchical clustering, *K*-means clustering, as well as *t*-tests were performed to identify specific trends in the gene expression during a WNV infection in human brain cells. Approximately 240 genes that were differentially regulated in at least 3 out of the 5 time points were used in these analyses. Genes that were differentially regulated in 2 or less of the time points were disregarded in this study, as they were not considered to be sufficiently significant.

### 4.3.1 Analyses using Hierarchical Clustering

The first method used was hierarchical clustering. Figure 4-6 shows an overview of the tree structure after hierarchical clustering was performed, while Figure 4-7 shows the enlarged top portion of the tree structure. A red cell denotes an upregulation of the

gene and a green cell denotes a downregulation. The intensity of the colour represents the magnitude of the fold-change, which can be compared with the colour bar at the top. A black cell denotes an insignificant change in expression and no numerical value is given to that gene. The tree node structures of the clustering can be observed at the top and left side of the figure, while the identities of the genes (GenBank accession) are listed on the right.

In this analysis, both the genes and experiments were both subjected to the hierarchical clustering algorithm. From the top of Figure 4-7, the experiments were observed to be clustered in the sequential time points. From the tree nodes, the algorithm indicated that the genes regulated in the first 3 time points (1.5 h, 6 h, and 12 h p.i.) were relatively different from the last 2 time points (18 h and 24 h p.i.). This is not coincidental as it represents the early/middle phase and the late phase of a virus infection, respectively. Clustering of the experimental time points had therefore helped to verify the validity of the microarray experiments.

Based on the expression values, the genes were clustered into 9 separate groups. The number of genes in each group and the average movements of the expression values over time are displayed in the centroid graphs (Figure 4-8). However, upon observation of the expression graphs in Figure 4-9, many of the genes within the clusters were found to be expressed in a rather inconsistent manner. Some of the genes did not follow the trend line (in pink) closely. Analyses of the genes in each cluster were therefore difficult as there was great variability in the gene expression values within the same cluster.

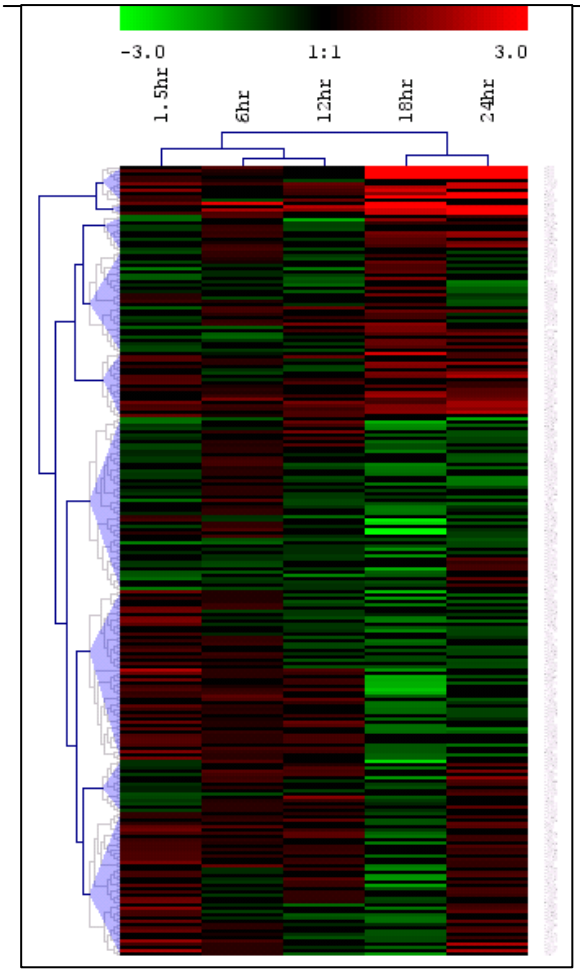
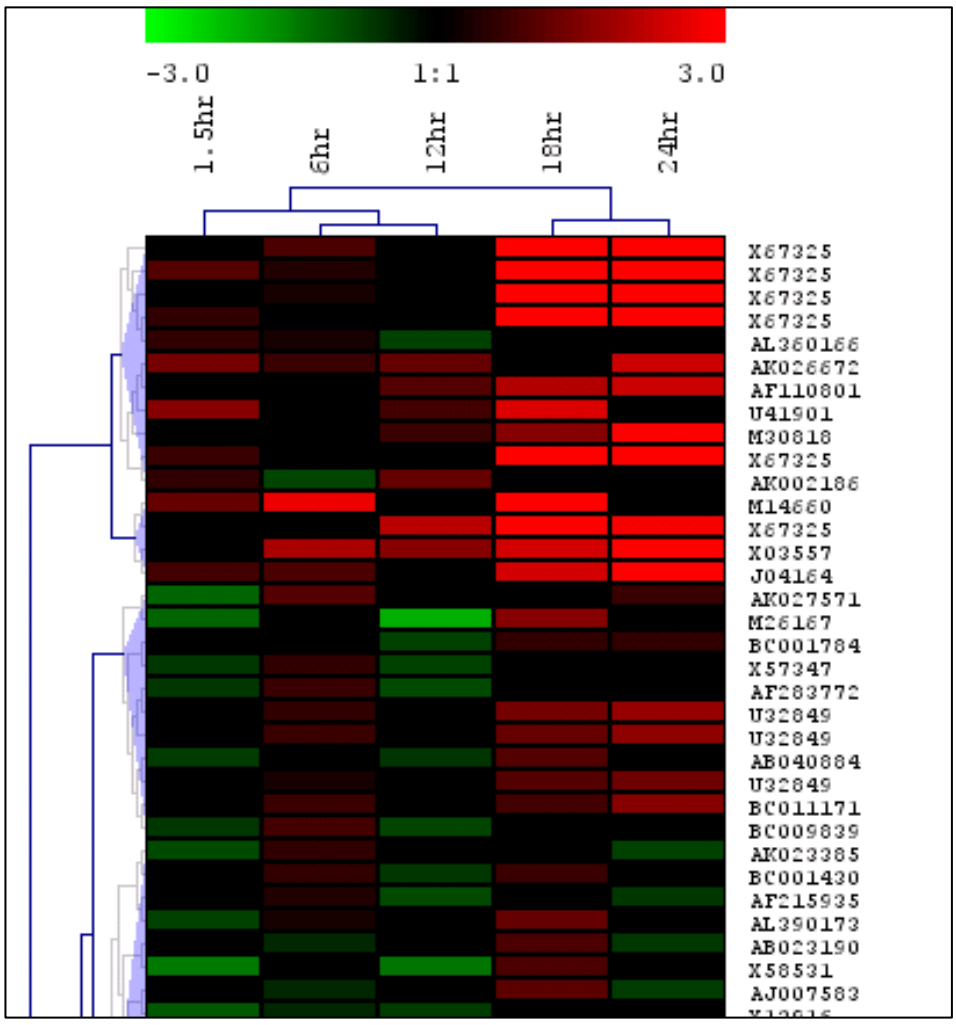



**Figure 4-6** (left). An overview of the tree structure from the hierarchical clustering analysis.

**Figure 4-7** (below). An expanded view of the first three node structures.



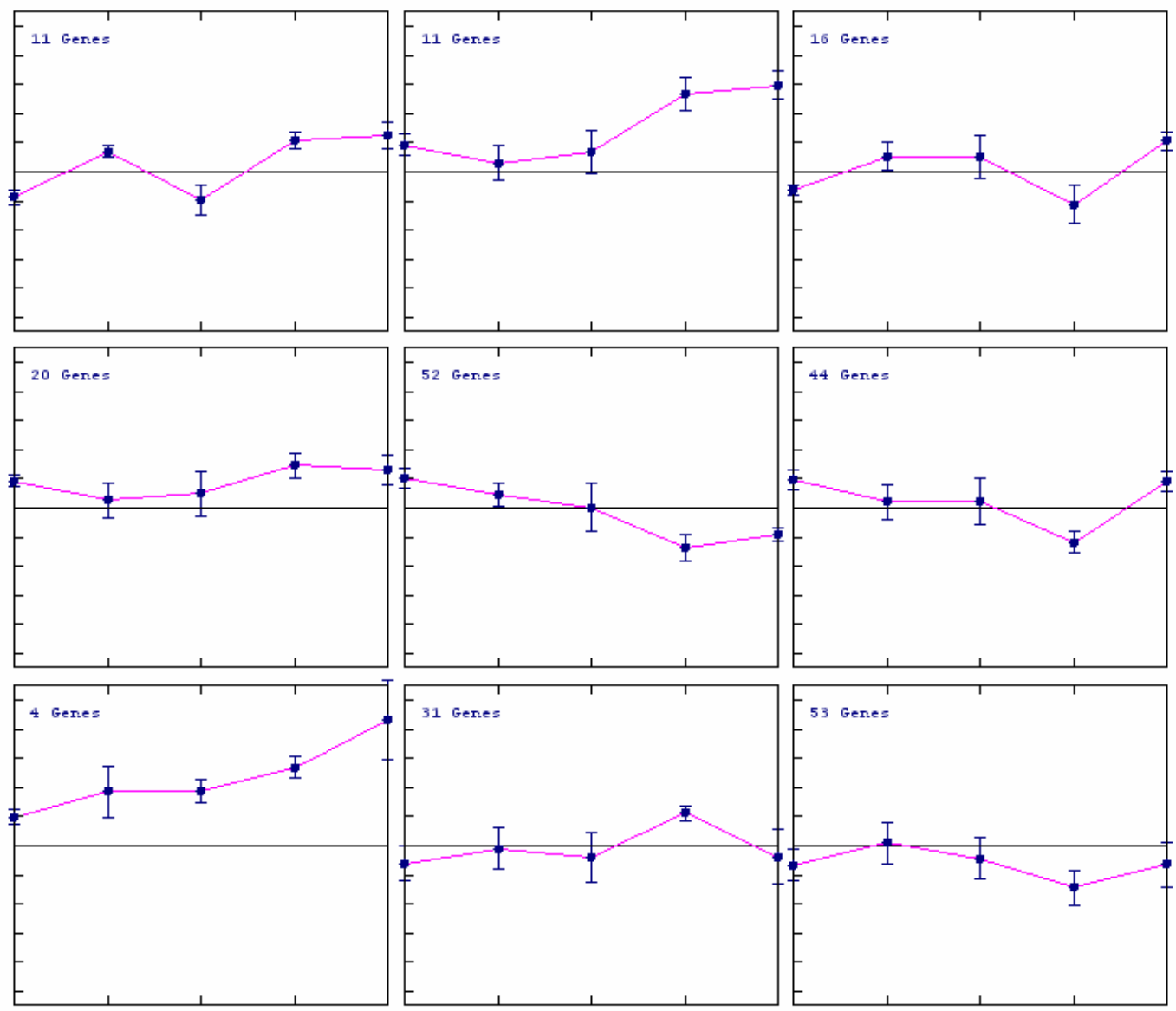



**Figure 4-8.** Centroid graphs of the 9 clusters from hierarchical clustering. The number of genes in each cluster is shown on the top left hand corner. The line shows the general trend of the gene expression pattern over the 5 time points in each cluster.

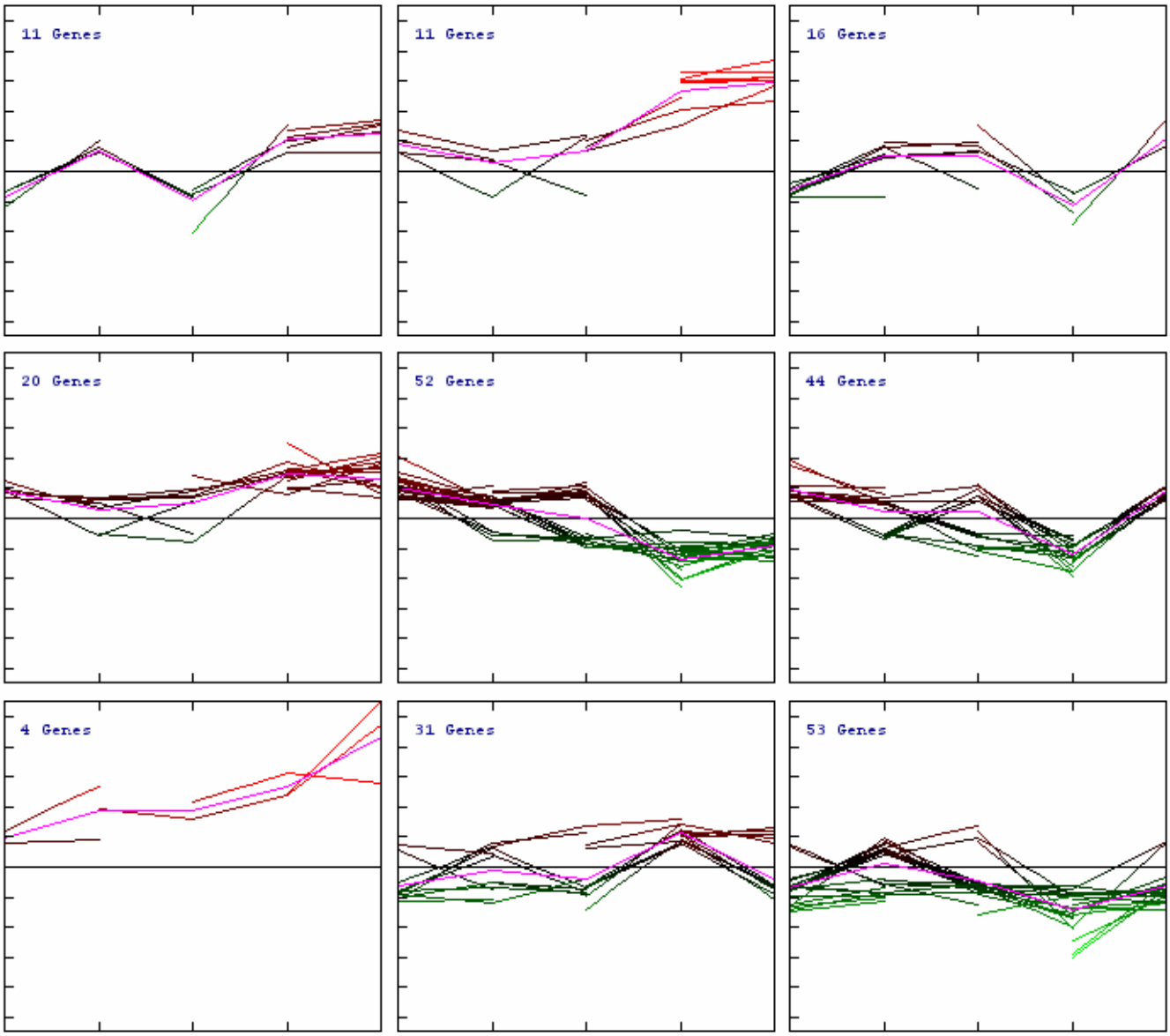



**Figure 4-9.** Expression graphs of the 9 clusters from hierarchical clustering. The individual lines show the expression pattern of each gene. Overall, the gene expression patterns clustered using hierarchical clustering is not consistent within each cluster, as many of them do not follow the trend line (in pink) closely.

### 4.3.2 Analyses using the Self-Organizing Tree Algorithm (SOTA)

The SOTA algorithm constructs a binary tree (dendrogram) in which the terminal nodes are the resulting clusters. From the results, 11 clusters were produced, and the general trend across these clusters is shown in the SOTA dendrogram (Figure 4-10). It displays the generated tree with the expression image of each resulting cluster's centroid gene. The text to the right of the centroid expression image includes the cluster number, the number of genes within the cluster, and the cluster diversity (or the mean gene to centroid distance). The centroid graphs and the expression graphs of the 11 clusters are shown in Figures 4-11 and 4-12, respectively.

When comparing the graphs with those obtained for hierarchical clustering, the clusters obtained from SOTA were less haphazardly arranged. The gene expression patterns in the SOTA clusters followed the centroid graphs rather closely, with the exception of clusters 1, 3 and 11 (figure 4-12). The better clustering in SOTA could be due to the creation of 2 extra clusters, and thus allowing the better fit in the grouping of genes.

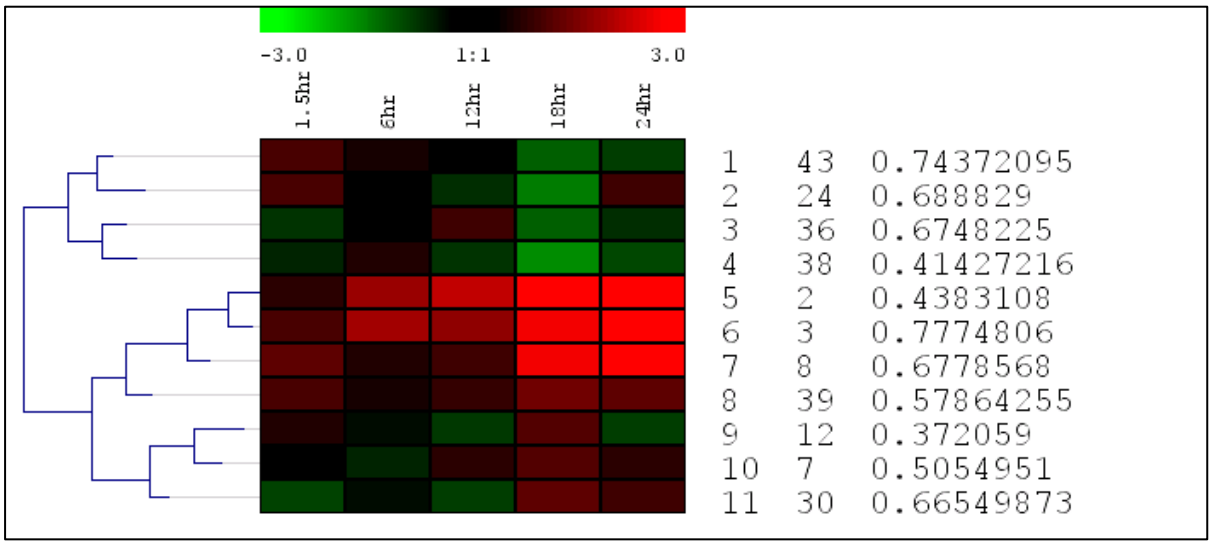



**Figure 4-10.** SOTA Dendrogram.

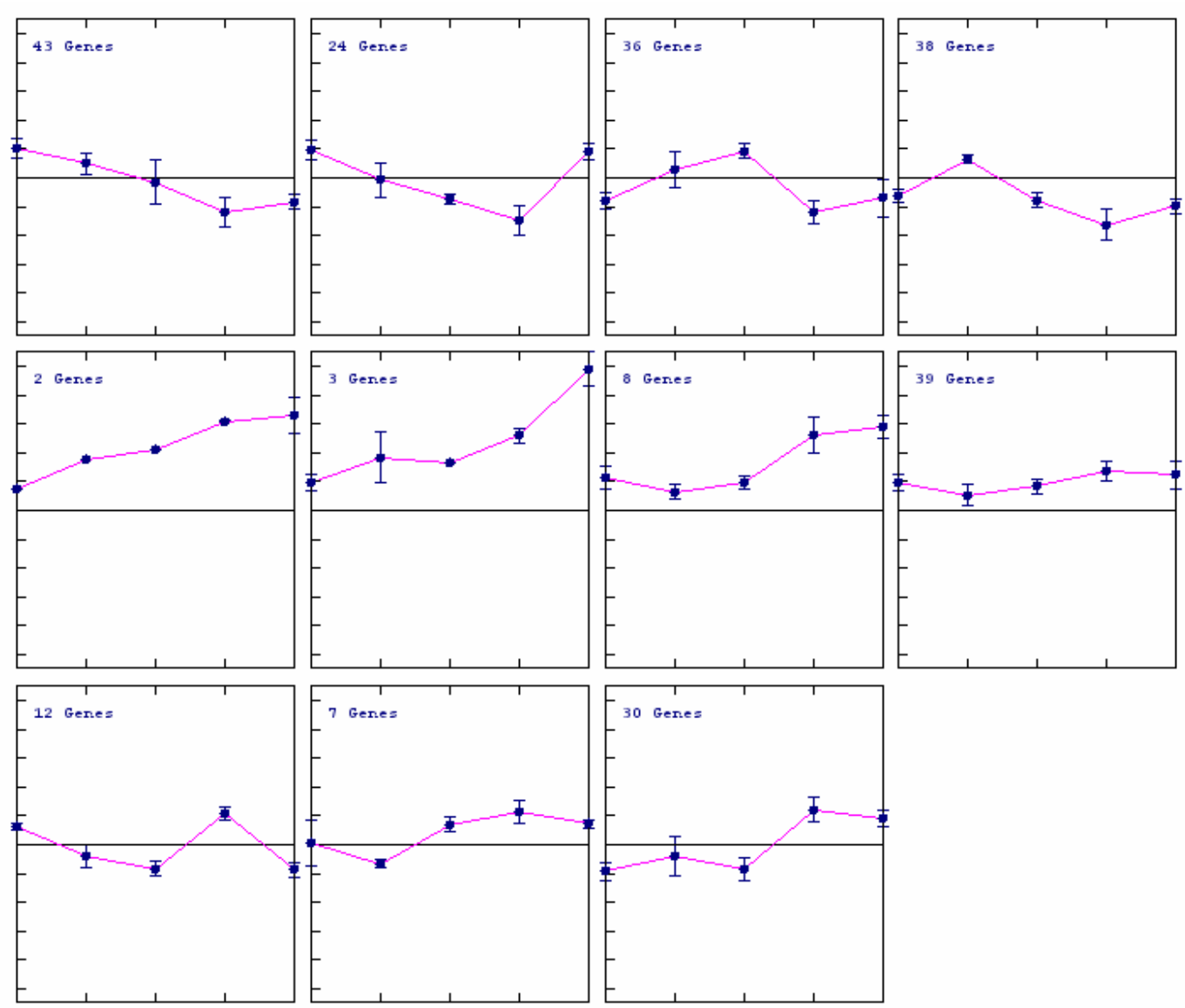



**Figure 4-11.** Centroid graphs of the 11 clusters from SOTA analysis.

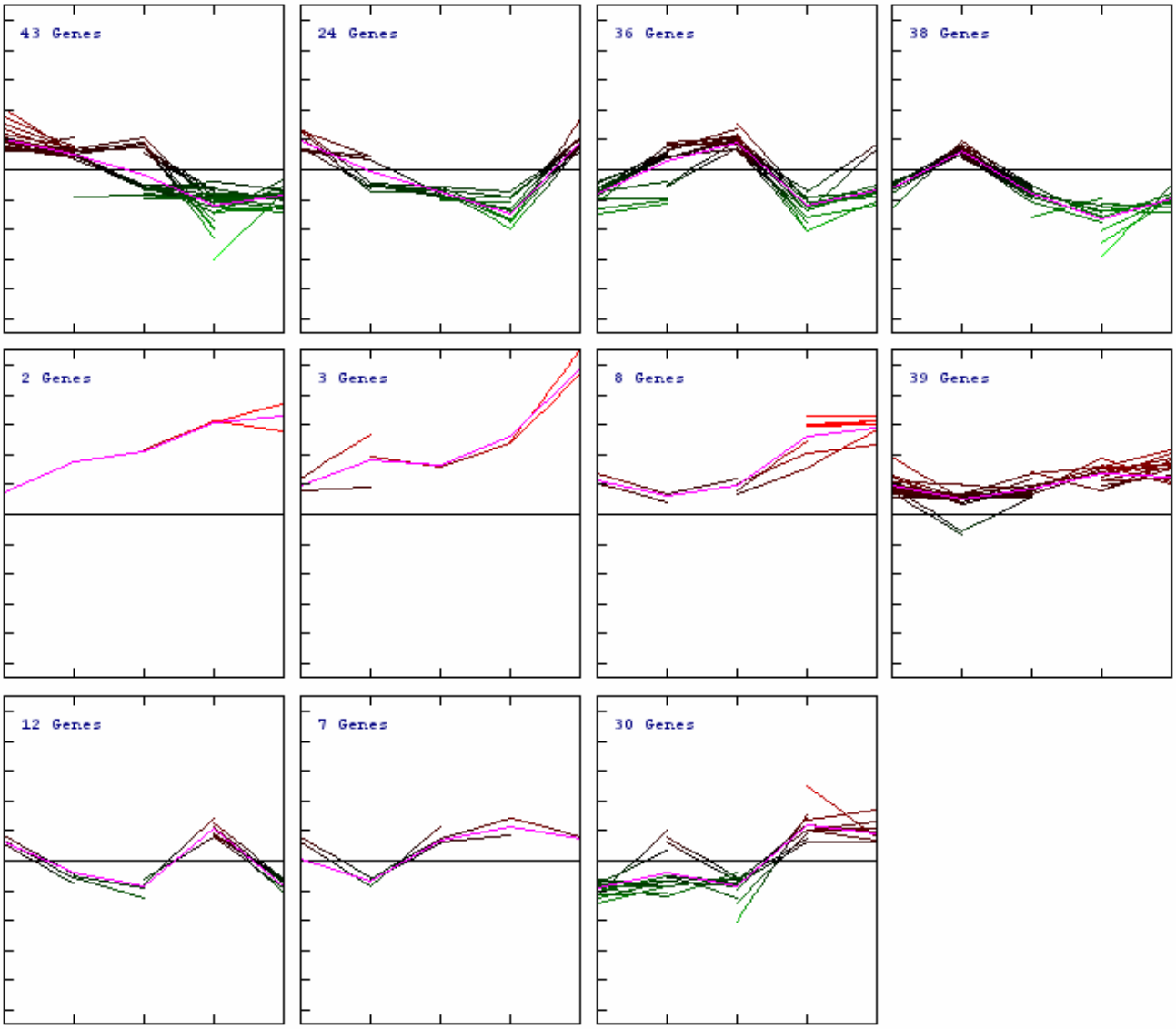



**Figure 4-12.** Expression graphs of the 11 clusters from SOTA analysis. The clusters are numbered sequentially from the top left to the bottom right. The genes in cluster 1, 3 and 11 do not seem to group together consistently, and do not follow the trend line (in pink) closely.

### 4.3.3 Analysis using *K*-Means Clustering

*K*-means clustering represents another popular algorithm often used in clustering genes in microarray experiments. In this case, the number of clusters can be specified and the algorithm will try to fit the genes into these groups. In order to determine the ideal number of clusters to be generated, another algorithm was used. The Figures of Merit (FOM) algorithm helps to measure the extent of fit for a group of genes in different number of clusters. The graph generated by FOM analysis is shown in Figure 4-13. As the number of clusters increases, the fit of the genes within the cluster becomes tighter and less haphazard, and therefore the value of the adjusted FOM decreases. A compromise must be obtained between the FOM value and the number of clusters, as too many clusters make analyses difficult. From the FOM graph, a good choice for the number of clusters would be either 8 or 10, while maintaining a manageable number of clusters.

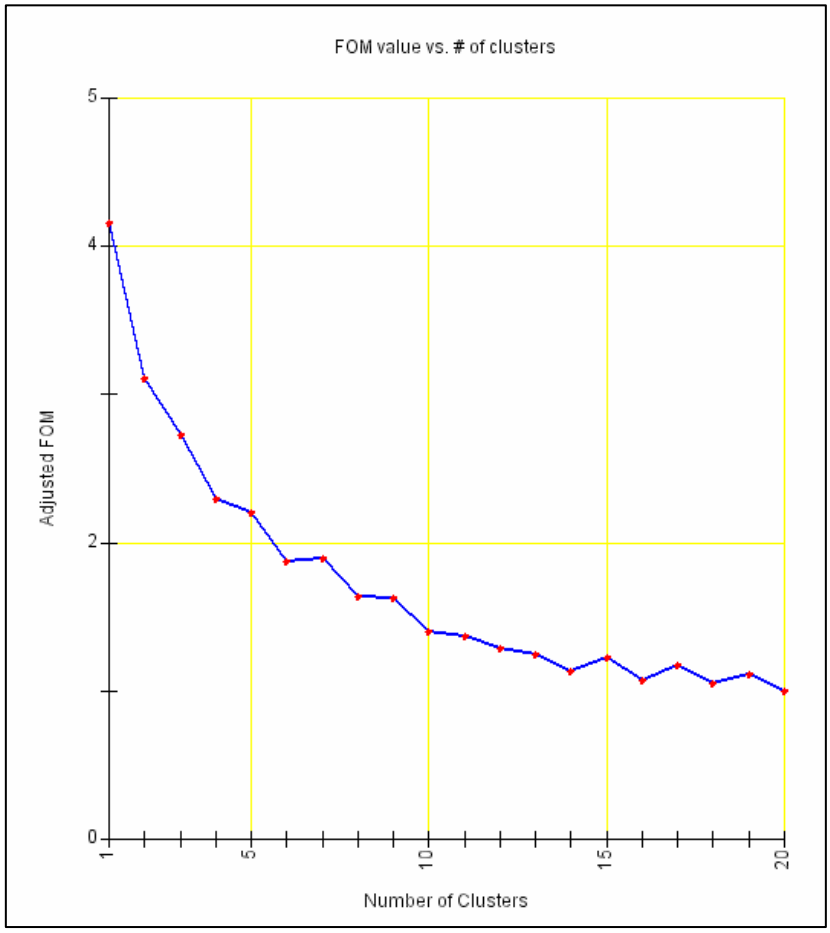



**Figure 4-13.** Figures of Merit (FOM) graph.

*K*-means clustering was therefore carried out by specifying the algorithm to generate either 8 or 10 clusters. The results of these analyses are shown in Figures 4-14 to 4-17. As can be seen from the graphs, *K*-means clustering algorithm appeared to produce better clustering results of the genes compared to hierarchical clustering and SOTA, with less of the haphazard arrangements seen in the previous methods. As the clustering of the genes appeared to be more coherent in the 10 clusters, subsequent analyses of gene regulation will utilize the data generated from that set. The disadvantage of using *K*-means is that the tree structures between the different clusters cannot be visualized (as compared to the clustering graphs in Figures 4-7 and 4-10). However, this is an insignificant consequence as the goal is to achieve reliable clustering of results.

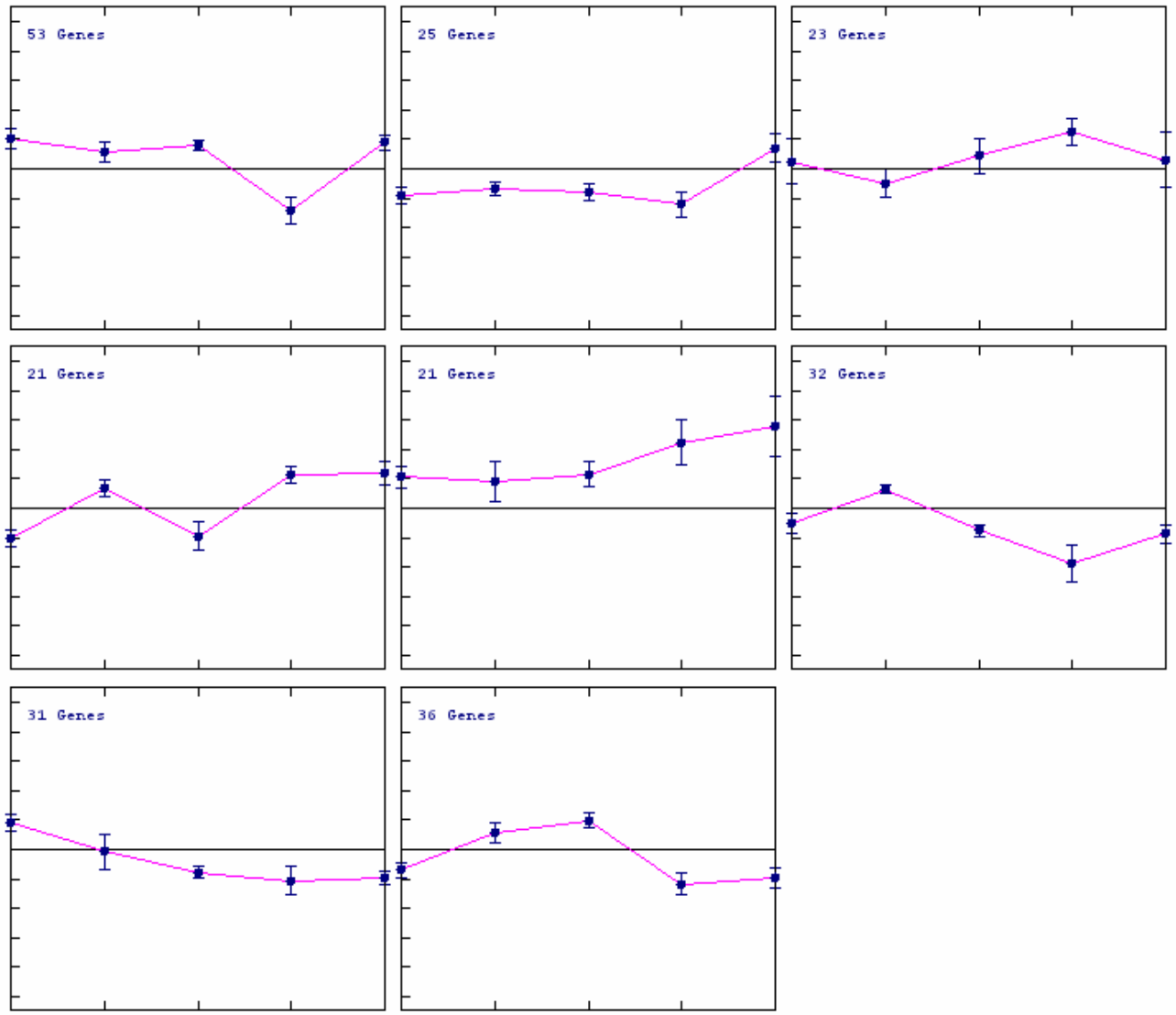



**Figure 4-14.** Centroid graphs of the 8 clusters from *K*-means clustering.

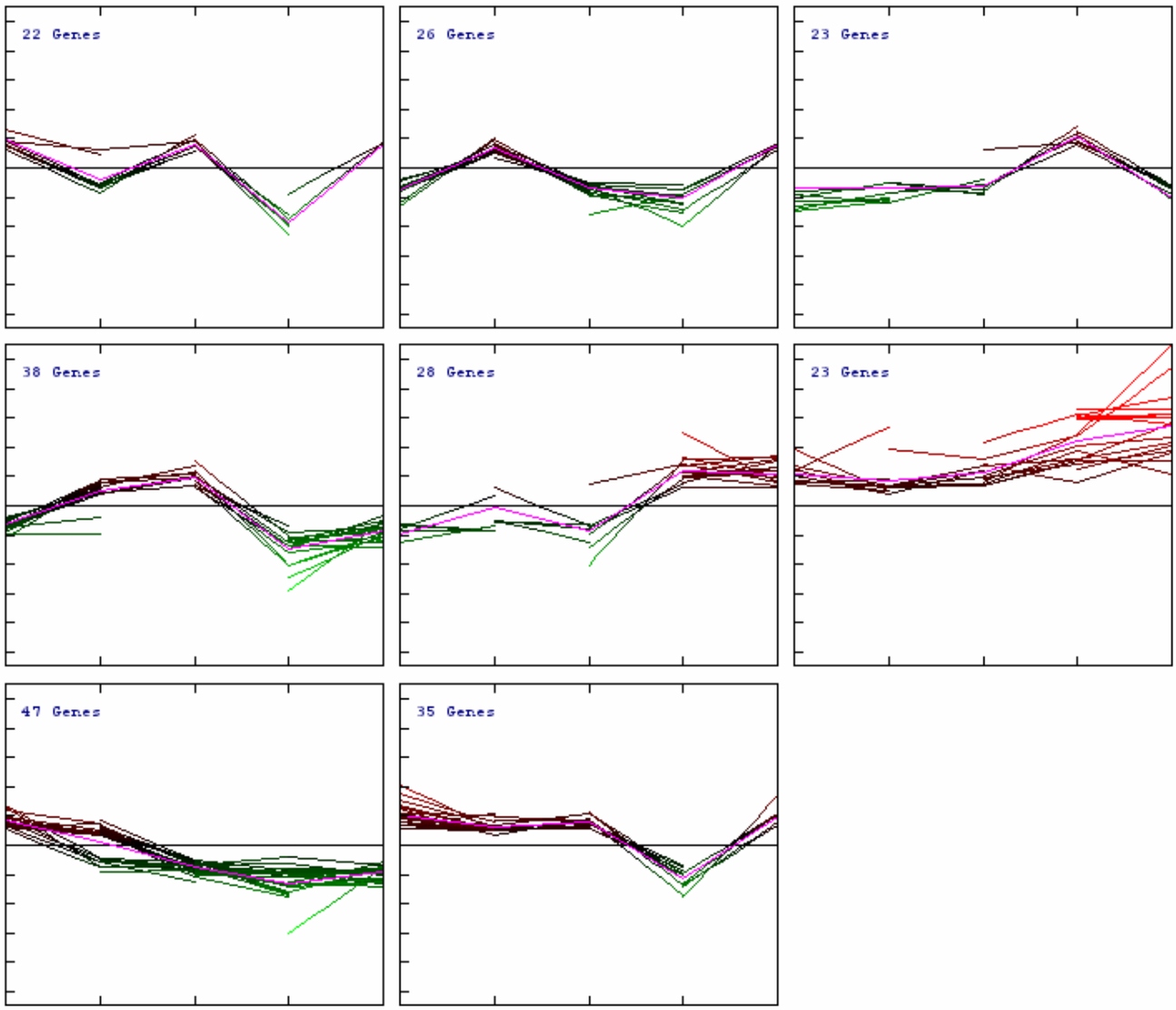



**Figure 4-15.** Expression graphs of the 8 clusters from *K*-means clustering.

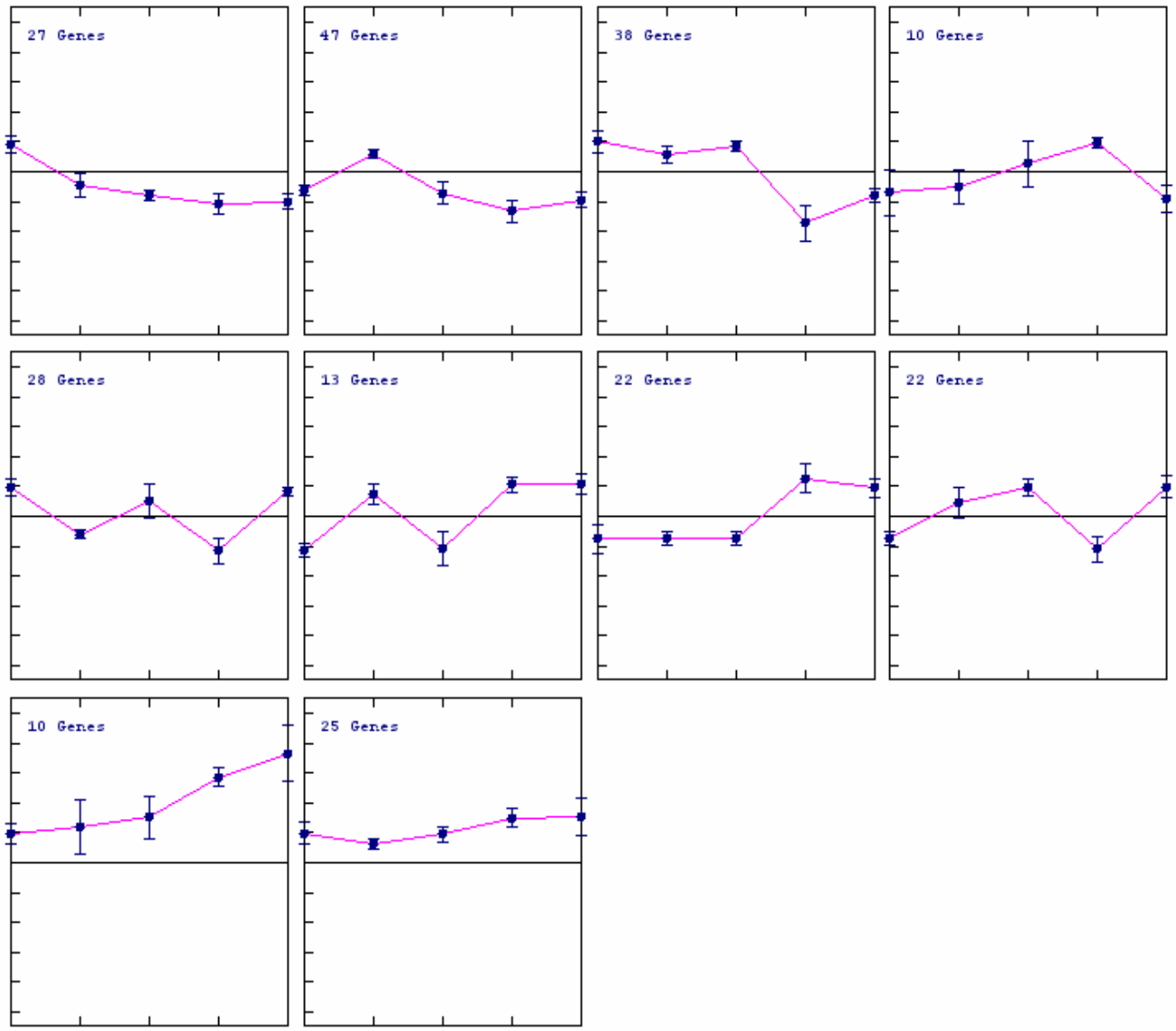



**Figure 4-16.** Centroid graphs of the 10 clusters from *K*-means clustering.

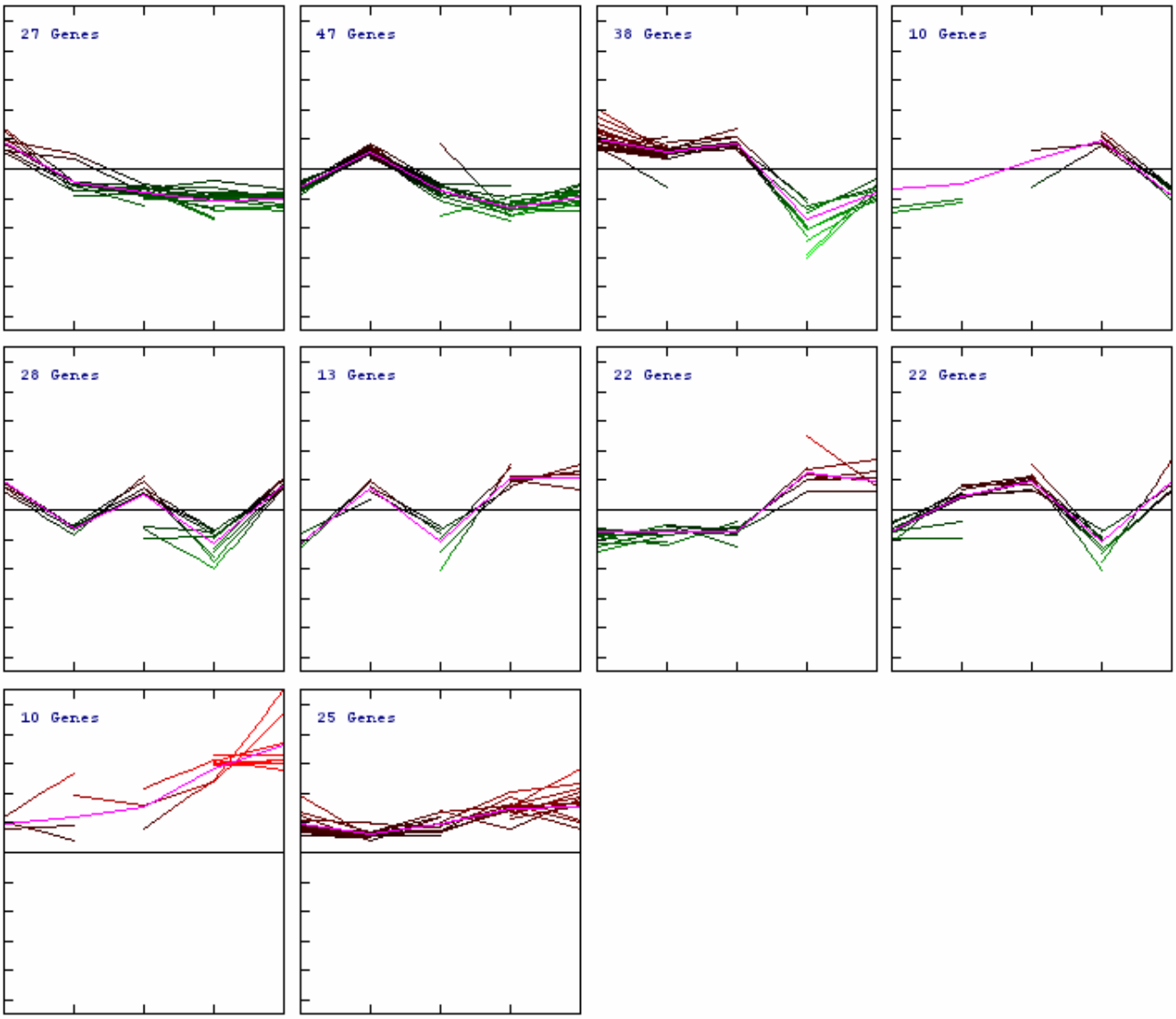



**Figure 4-17.** Expression graphs of the 10 clusters from *K*-means clustering.

### 4.3.4 Analyses using T-Test Statistics

By using the statistical one-sample *t*-test to analyze the results, genes which are statistically different from a mean value of zero expression at $p$ <0.05, will be clustered into one group. This will identify genes which were either continually upregulated or downregulated throughout the study. Forty-six genes were identified in this process, and the hierarchical tree structure is shown in Figure 4-18. The expression graph of the same set of genes is shown in Figure 4-19. As can be seen in the hierarchical tree, these 46 genes can be divided into 2 main subgroups. Those in red are consistently upregulated through the time series, whereas those in green (the second node) are consistently downregulated. The first 5 genes can be further subdivided from the rest of the upregulated genes, as their expression at 18 h and 24 h were exceedingly high.

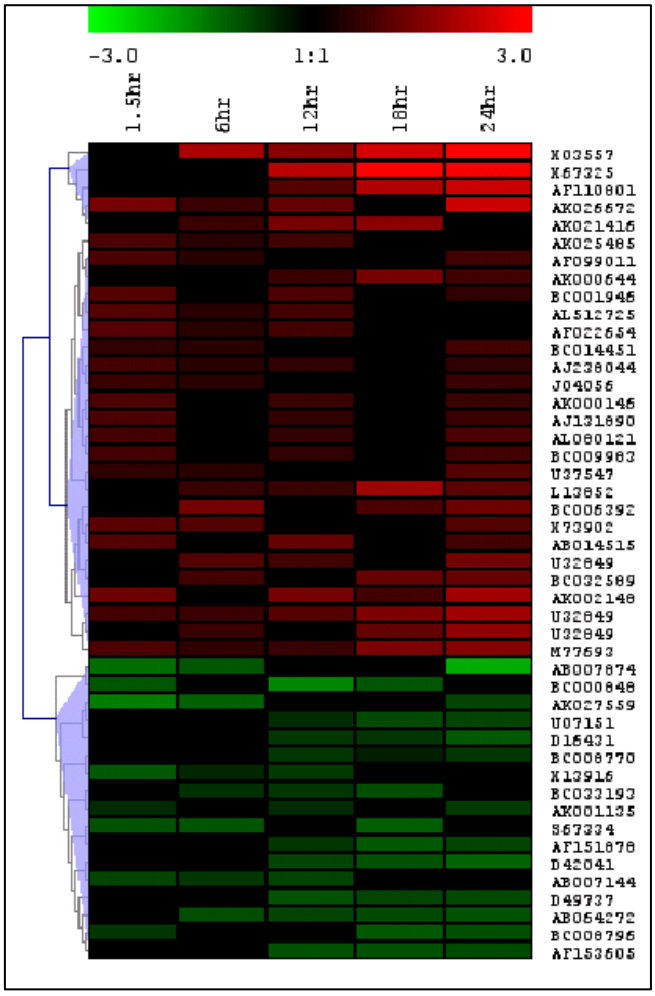



**Figure 4-18.** Hierarchical tree of statistically significant genes from the *t*-test.

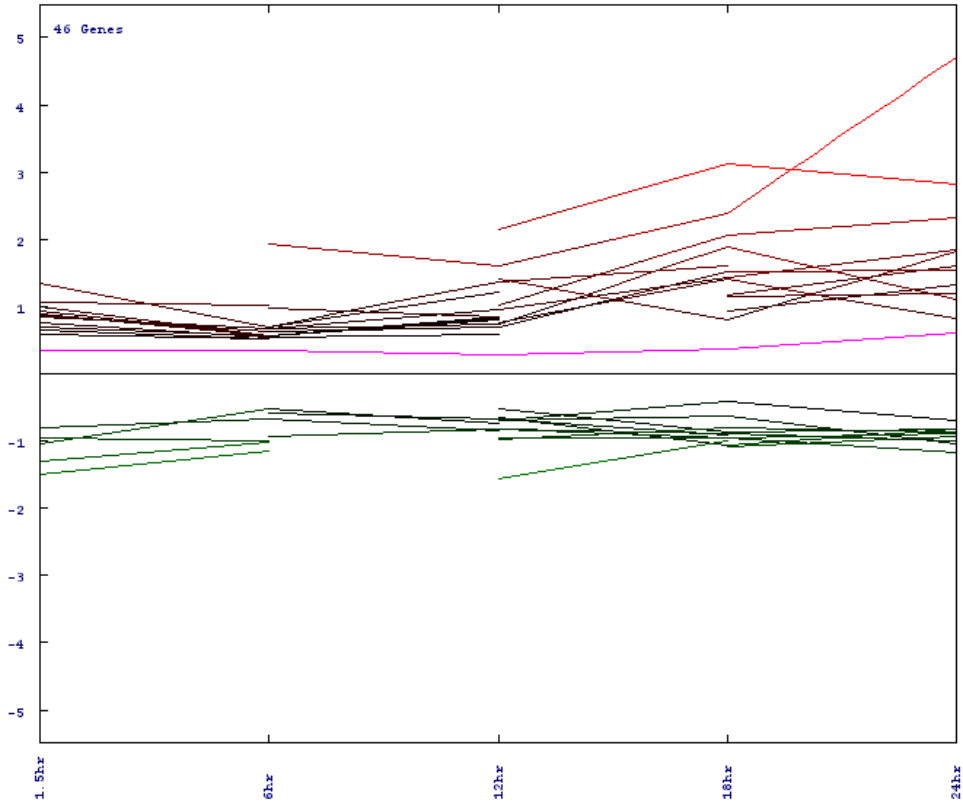



**Figure 4-19.** Expression graph of significant genes from *t*-test.

## 4.4 Identifying Trends in Gene Expression

From the centroid and expression graphs obtained from the various clustering algorithms, certain gene expression trends during the course of the virus infection can be identified.

The most obvious group of genes was those which were upregulated for all the 5 time-points during the study. This group of genes includes those found in SOTA (clusters #5, 6, 7 and 8), *K*-means 8-clustering (cluster #6), and *K*-means 10-clustering (clusters #9 and 10). A representative expression graph with genes that are consistently upregulated is shown in Figure 4-20. The genes from all these clusters were combined, and repetitive genes present in 2 or more clusters were removed. Forty-five genes were identified using this method, and these are listed in Appendix 7. This group will subsequently be referred to as GpI.

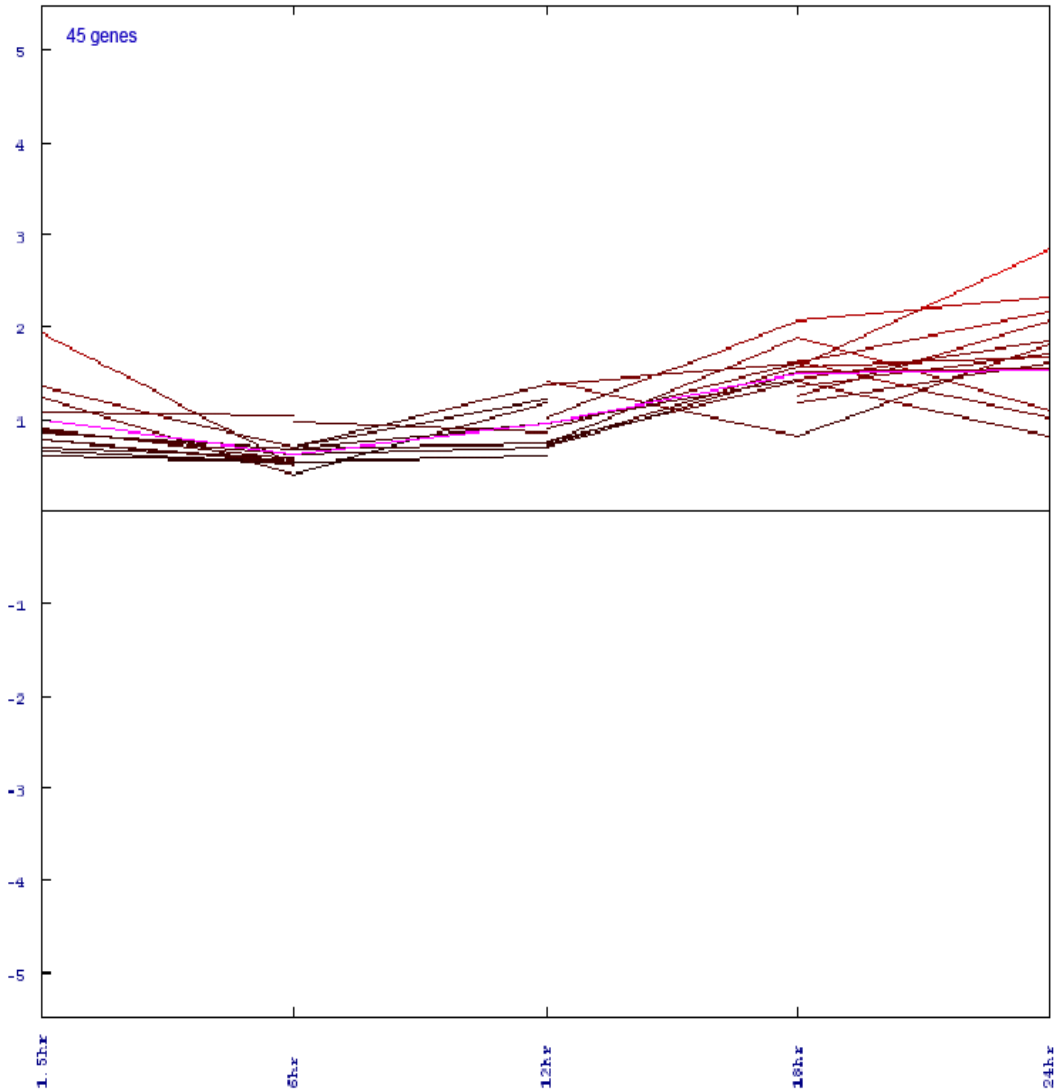



**Figure 4-20.** An expression graph of constantly upregulated genes throughout the course of WNV infection. (GpI).

To identify the ontology group that was over represented in this upregulated cluster, these 45 genes were loaded into *EASE*. The immune response functional group (15 genes) was highly significant with a *Fisher's Exact* value of 7.38E-10, and these are highlighted in yellow.

On the other hand, genes which were constantly downregulated were much fewer. Only the *t*-test analysis was able to cluster this group of genes (Figure 4-19). These 17

genes are listed in Appendix 8, and are labelled as GpII. When these were loaded into *EASE*, no significant ontology group was found with a *Fisher's Exact* value of < 0.05. These constantly downregulated genes were thus quite diverse in functions. Figure 4-21 shows an expression graph for GpII.

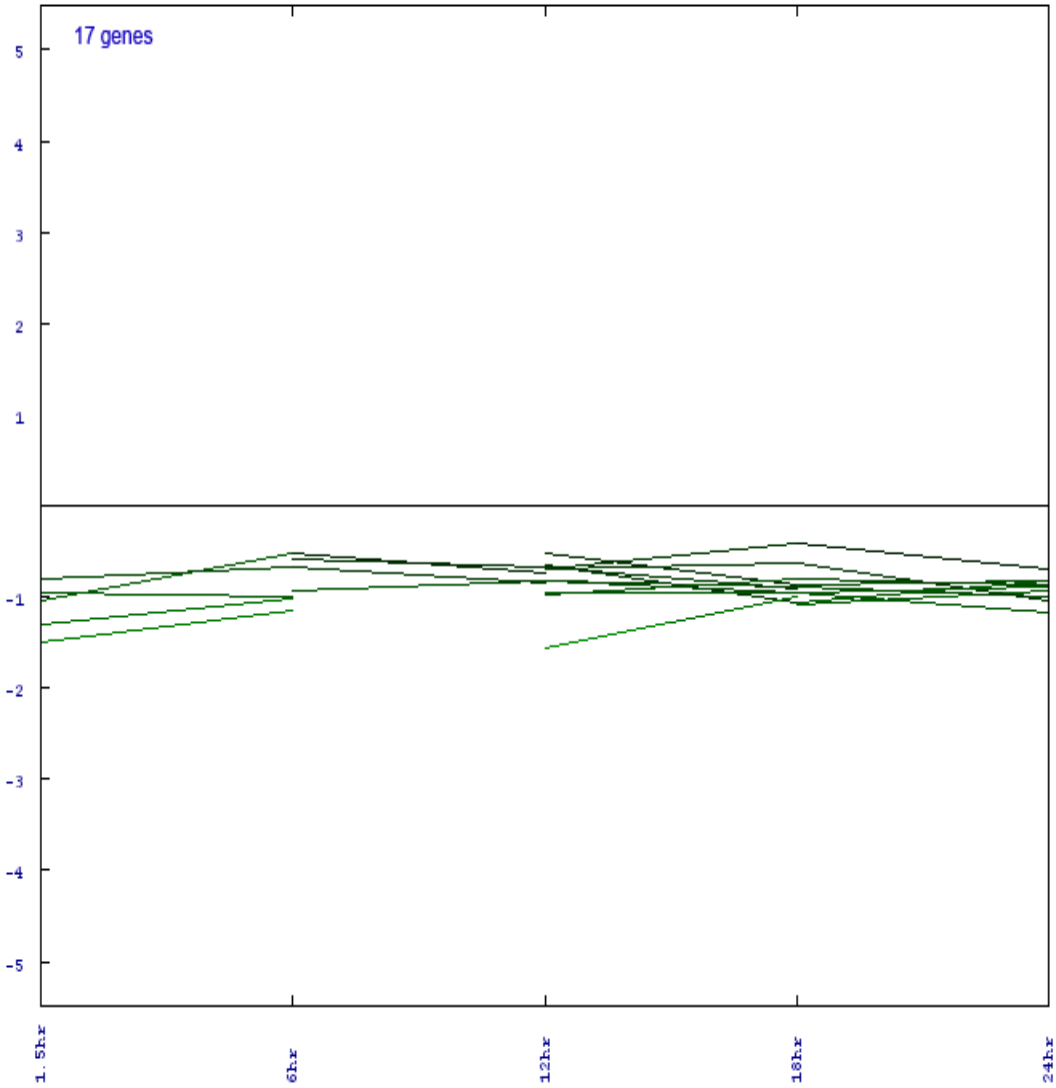



**Figure 4-21.** An expression graph of genes that are constantly downregulated throughout the course of an infection. (GpII)

The next group of genes to explore was those which were upregulated initially, but were later downregulated as the course of infection progresses. Two clusters from *K*-means clustering fit this profile. The first group was downregulated just after 6h p.i.

(GpIII), while the second group was downregulated after 18h p.i. (GpIV). Their expression profiles are shown in Figures 4-22 and 4-23 respectively.

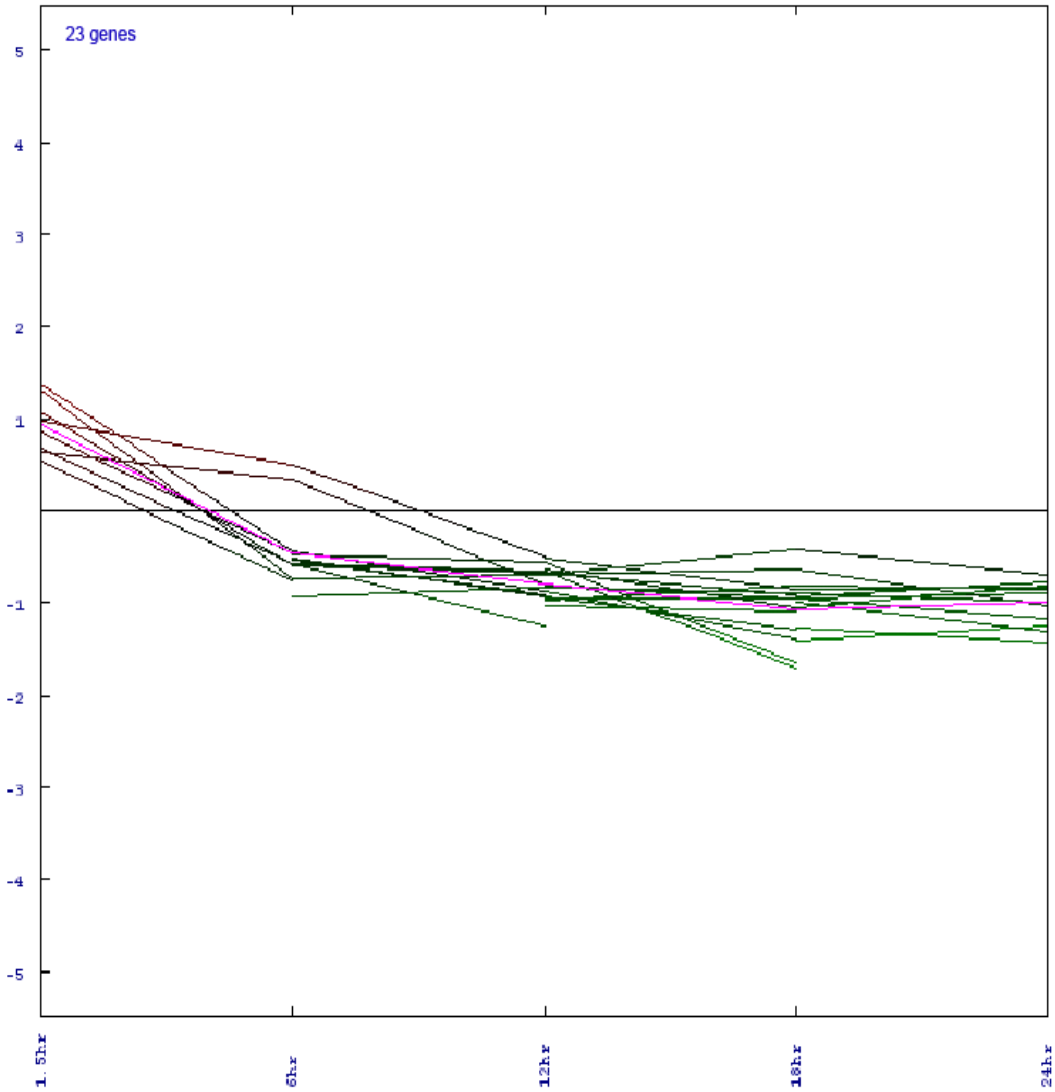



**Figure 4-22.** An expression graph of genes that were initially upregulated, but were downregulated after 6 h p.i. (GpIII)

Twenty-three genes were clustered into GpIII. The ontology group that was found to be over-represented by *EASE* belongs to the cell cycle regulation group with a *Fisher's Exact* value of 0.014. These genes are highlighted in yellow in Appendix 9. Thirty-eight genes were clustered into GpIV. Genes belonging to transcription and nucleic acid metabolism ontology group were found to be significantly over-represented with a *Fisher's Exact* value of 0.0282. These genes are highlighted in yellow in Appendix 10.

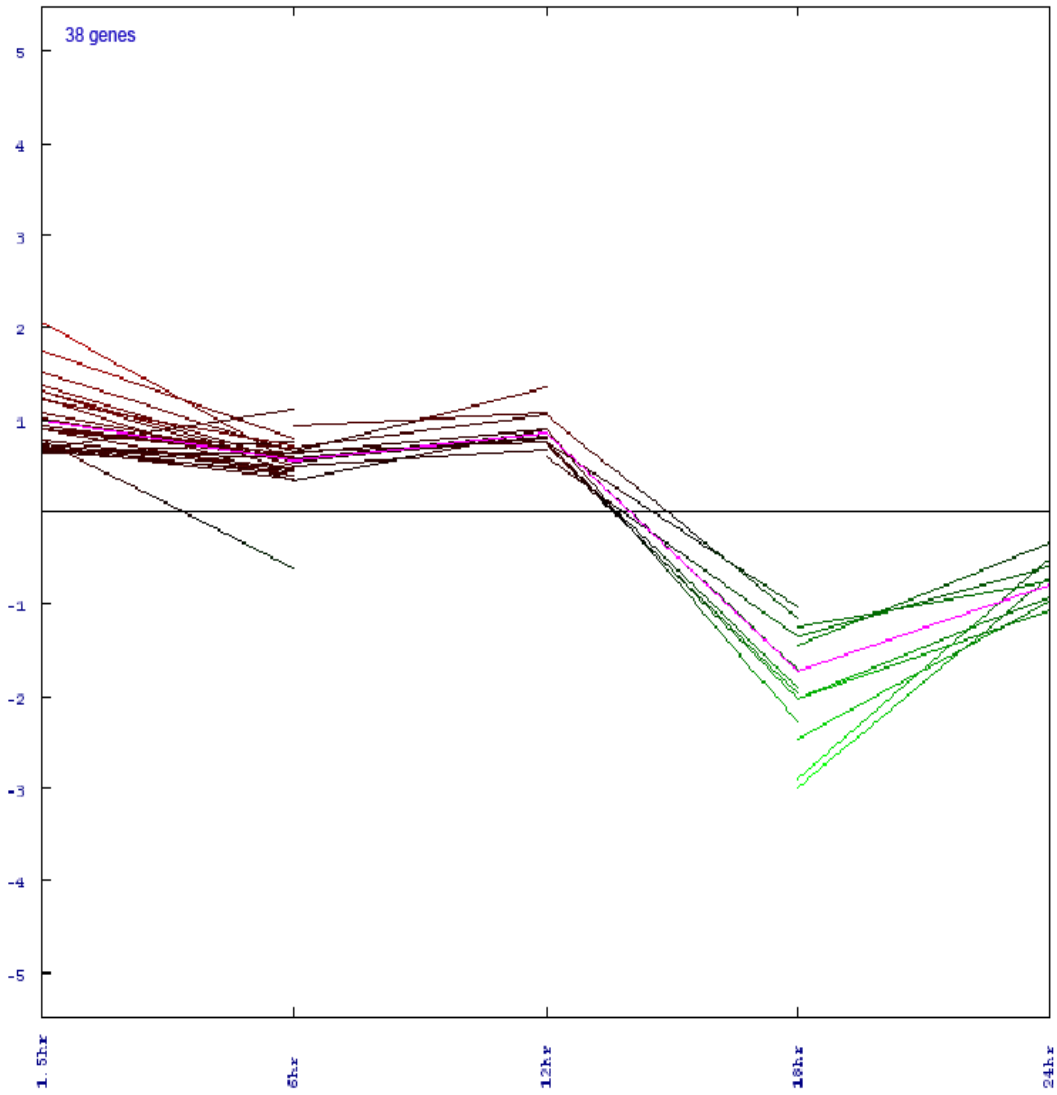



**Figure 4-23.** An expression graph of genes that were initially upregulated, but were downregulated after 18 h p.i. (GpIV)

The next group of genes was those whose expression levels increased as the infection progressed, and this cluster is labelled as GpV. Clusters belonging to this group are represented in Figure 4-24, and they were obtained from SOTA (cluster #11) and K-means (cluster #7). Interestingly, these genes were downregulated during the initial phase of infection, but became upregulated after 18h p.i.

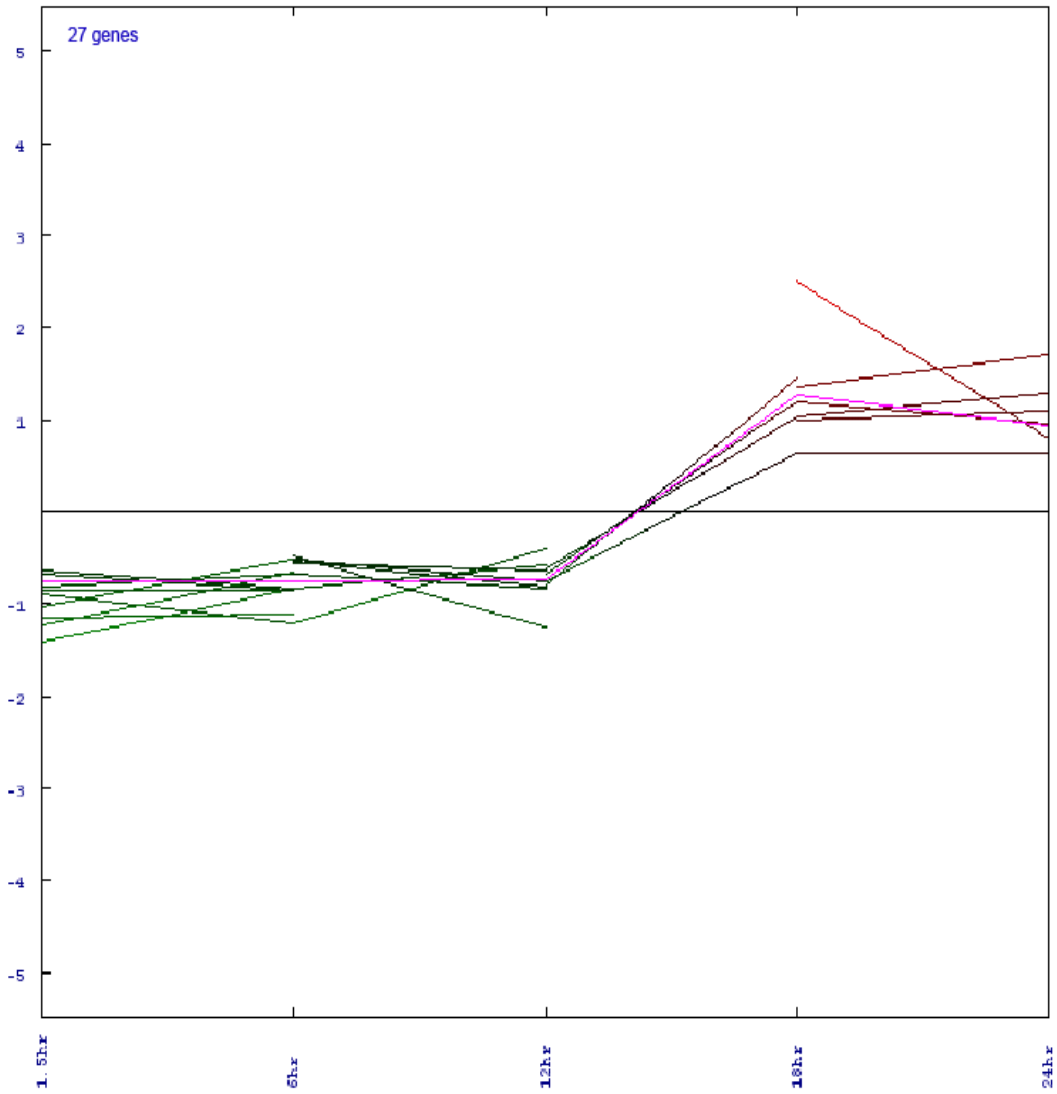



**Figure 4-24.** An expression graph of genes that were initially downregulated, but were upregulated after 18 h p.i. (GpV) The expression pattern of this group is the exact opposite of Gp IV.

Twenty-seven genes were represented in GpV. Genes belonging to the cellular physiological processes and proteins integral to plasma membrane ontology groups were found to be significantly over-represented with a *Fisher's Exact* value of 0.0224 and 0.0263, respectively. These genes are highlighted in yellow in Appendix 11.

The last group of genes shows a peculiar pattern of expression. These genes were initially downregulated at 1.5 h p.i., but became upregulated at 6 h, and finally returned to be downregulated for the next three time-points (12 h, 18 h, and 24 h). Genes whose expression patterns belonging to this group (GpVI) are represented in Figure 4-25, and they are obtained from SOTA (cluster #4) and K-means (cluster #2).

Fifty-one genes were represented in GpVI. Three ontology groups were found to be significantly enriched in this group, and these genes belong to the endoplasmic reticulum, mitochondrion and cytoskeletal protein binding functional groups. These genes are also highlighted in yellow in Appendix 12.

A summary of the 6 groups from the microarray results is listed in Table 4-4.

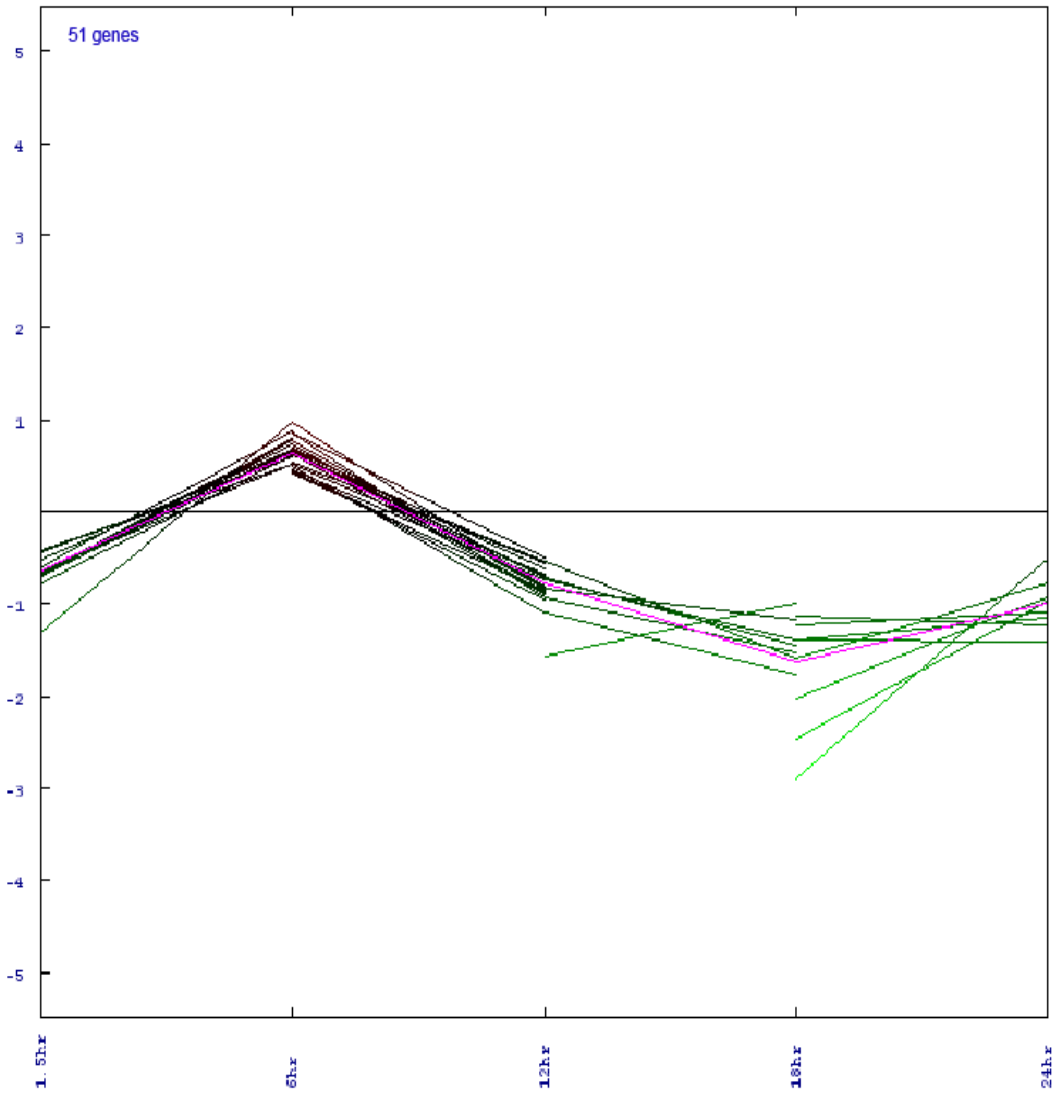



**Figure 4-25.** An expression graph of genes that were consistently downregulated, except at 6 h p.i. (GpVI)

**Table 4-4.** Summary of the 6 groups from microarray analysis.

| Groups | Gene Expression Trend | Number of Genes | Significant Gene Ontologies |
|---|---|---|---|
| **I** | Constantly upregulated | 45 | Immune response |
| **II** | Constantly downregulated | 17 | None |
| **III** | Downregulation after 6hr of infection | 23 | Regulation of cell cycle |
| **IV** | Downregulation after 18hr of infection | 38 | DNA-dependent transcription<br>Nucleic acid metabolism |
| **V** | Upregulation after 18hr of infection | 27 | Cellular physiological processes<br>Integral to plasma membrane |
| **VI** | Upregulation only at 6hr | 51 | Endoplasmic reticulum<br>Mitochondria<br>Cytoskeletal functions |

# 5.0 DISCUSSION

## 5.1 Cytopathic Effects of West Nile (Sarafend) Virus Infection

West Nile (Sarafend) virus [WN(S)V] infection in HeLa (cervical adenocarcinoma) and A172 (glioblastoma) cells resulted in two outcomes. Although cytopathic effects (CPE, Figure 3-2) can be observed in HeLa cells, they were not extensive and they produced lower virus yields. This was confirmed by studies using plaque assays, quantitative PCR, and immunofluorescence microscopy. Dunster and colleagues (1990) had also reported similar findings. It was found that the WNV became attenuated in just six passage cycles, with a 4 X $10^3$ fold reduction in virus yield. It can therefore be deduced that HeLa cells are not readily permissive to WNV infection. Its less permissive nature may lead to novel inhibitory mechanisms to WNV infections.

In contrast, WNV infection in A172 (glial) cells showed advanced CPE within 24hr (Figure 3-4). Plaque assay studies showed a 100-fold increase in virus yields from infected-A172 cells compared to infected-HeLa cells (Section 3.3). The intracellular levels of the virus was also about 4-times more in A172 cells, as determined by immunofluorescence and quantitative PCR studies (Section 3.4 and 3.5). This was not surprising as WNV is neurotrophic, thereby causing encephalitis. Neuronal degeneration, neuronophagia, spongy degeneration, and focal haemorrhages have been observed in the thalamus, cerebrum, cerebellum, medulla oblongata and cervical spinal cord of humans infected with the WNV (Sampson *et al.*, 2000). The WNV cellular receptor αVβ3 integrin, was found to be expressed in equal amounts in both

A172 and HeLa cells (Chu and Ng, in press), and therefore virus entry should not be the limiting factor.

This study was therefore conducted in two parts to investigate in detail the differential host responses during WNV infection. The first part of the study was to investigate the mechanisms for an apparent attenuation in virus replication in HeLa cells. West Nile (Sarafend) virus was allowed to infect the A172 and HeLa cell samples for 24 hours (see Section 2.2.2), after which, the cells were harvested for subsequent analysis. Twenty-four hours was chosen to ensure that the virus had undergone a full replication cycle and represented peak virus production. As A172 cells were found to be more permissive for WNV infection, the second part of the study was to investigate the host cell response at various time points (1.5 h, 6 h, 12 h, 18 h, and 24 h p.i.). This will allow the study of virus-host interactions at various stages of the virus replication.

## 5.2 Global Transcriptomic Analysis using Microarrays

Genomic expression studies on host-virus interactions using microarrays have the advantage of producing a huge volume of data (6MB of raw data per slide) in a single experiment. This mean that time spent in the lab per data point collected has become very small. However, a standard approach to analyze microarray data is still absent as methods in experimental techniques and in data analysis are continually being improved.

Microarray analysis is a complex system and is affected by a number of experimental factors, including target preparation, physical deposition of the targets, slide

chemistry, probe chemistry, hybridization, and detection of the fluorescent signal (Worley, 2000). Problems inherent to the technologies have raised important issues on how to apply adequate statistical tests. As a consequence, statistical approaches to microarray research in the life sciences are not yet as routine as they are in other sciences (Nadon and Shoemaker, 2002). The difficulties stem primarily from the myriad of potential sources of random and systematic error in the microarray process, and the absence of a systematic method for data analysis. Indeed, the same analyses performed using BRB ArrayTools and TIGR's MeV can produce different sets of results. And as also shown in the Results Section 4.3, utilization of different clustering algorithms can produce a range of different results.

In an attempt to address these issues, microarray experiments in this study was carried out in a meticulous manner. Three types of quality control values were used to verify data quality: (1) A series of metrics reported by the image analysis software to ensure that the spots that were quantified appeared to be good spots, (2) A series of controls on the microarray to ensure that hybridization had occurred correctly. These will indicate how specific and efficient the hybridization had been, (3) Analysis of the data from duplicate targets to indicate how good the overall data was and whether the results obtained were statistically meaningful (Amersham, 2002).

A lower limit of detection threshold must be set for each scanned microarray, such that spots with intensities below the lower limit are discarded. For this study, a more stringent lower limit was utilized. The lower limit was defined as three standard deviations of the pixel intensities on the negative control spots. This value was higher than the mean pixel intensities of the negative control spots, and thus, led to higher

stringency. The raw data generated in the Appendices therefore show a higher confidence of accuracy.

Direct incorporation of Cy3 or Cy5 modified nucleotide analogues have been known to introduce a gene or sequence-specific artefact (Tseng *et al.*, 2001). This difference in gene-specific incorporation efficiency manifests itself in a way that some genes produce consistently larger signals when labelled with one dye compared to the other. It has also been reported that Cy3- or Cy5- conjugated nucleotides are bulky, which makes their incorporation using standard enzymes very inefficient. In addition, rates of incorporation can differ between dyes, potentially resulting in dye biases (Holloway *et al.*, 2002). This type of experimental artefact can be most efficiently identified and flagged by exchanging dyes on samples (Bilban *et al.*, 2002). This procedure was carried out in this study.

Even for the normalization step, there are variations in how this can be carried out. Linear normalization of the data, where all the data points in the microarray are adjusted to one common value, is usually carried out. However some bioinformaticians had disputed that this was not accurate, and non-linear normalization should be carried out instead (Tseng *et al.*, 2001). Another method is to normalize against a putative housekeeping gene or a group of such genes (Hess *et al.*, 2001). There are a few other methods for normalization. Bilban and colleagues (2002) and Quackenbush (2002) have written exhaustive articles on the various aspects of normalizing microarray data. It is noteworthy to point out that since there are various ways to carry out normalization, there is perhaps no best method for normalization.

However, different methods of normalization are very likely to produce different set of results. Hoffmann and colleagues (2002) employed four different normalization methods for detection of differentially expressed genes on a prototype dataset. They found that the number of genes detected as differentially expressed differs by a factor of about three. This means that some genes that are truly differentially regulated will escape detection, and false positives will end up being detected. The dataset produced in this study must therefore be rigorously tested in a statistical manner. For this study, non-linear normalization using the Lowess algorithm was adopted (Yang *et al.*, 2002b), as it is the most commonly used and accepted form of normalization.

Microarray studies are often not an end to itself, as there is a need to validate the large data sets that are generated (Chuaqui *et al.*, 2002). Molecular 'gold-standard' techniques such as qRT-PCR or northern blots are often used to validate the results (Rajeevan *et al.*, 2001). These are termed lab-based confirmation of microarray data. Another method of validation is the *in silico* analysis where array results are compared

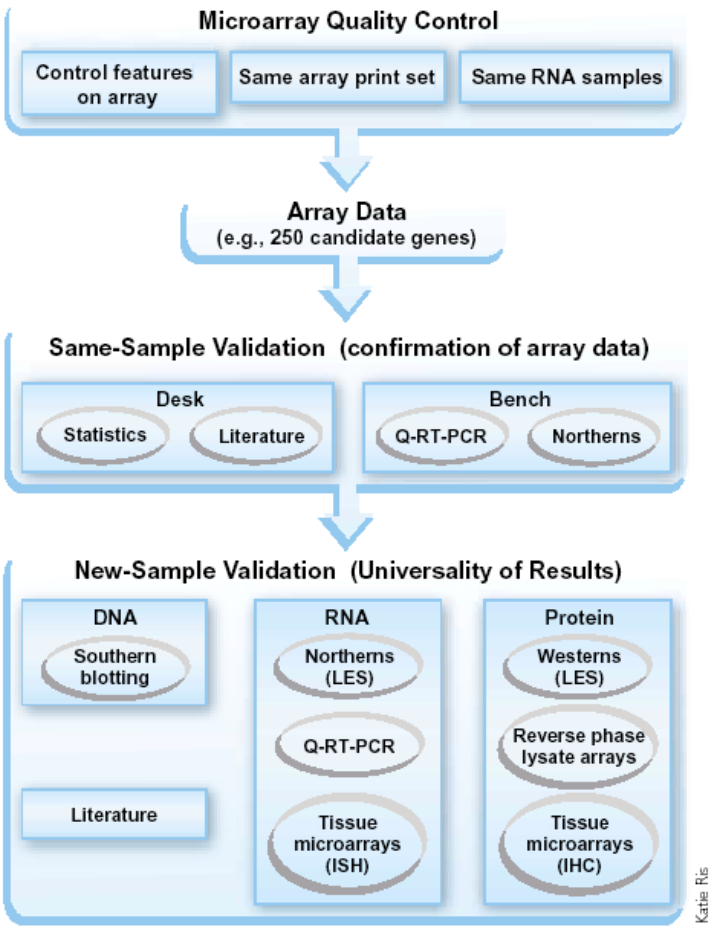



**Fig 4-1.** Key issues for validation of microarray data. The top panel lists the front-end quality control needed to facilitate successful downstream validation of microarray experiment results. The middle panel lists the minimum statistical steps required before follow-up studies. The bottom panel diagrams validation methods currently used to evaluating array results. These methods are used to address both independent confirmation of data and universality of the results (Chuaqui *et al.*, 2002).

with information available in the literature and in public or private expression databases. This provides the opportunity to validate data without further experimentation (Chuaqui *et al.*, 2002). Figure 4-1 shows the various techniques for microarray data validation. Both these methods of validation were carried when analyzing the microarray results. The *in silico* method of validation is particularly important in the face of a myriad of genes to analyse. However, due to the high number of genes involved in this study, it was not possible to carry out an exhaustive analysis of each gene individually. Some genes whose functions have yet to be characterized also served to limit the understanding of host-virus interactions.

## 5.3 Global Transcriptomic Comparison between HeLa and A172 Cells

### 5.3.1 Aberrations in Host Response in A172 Cells Lead to Observed Cytopathology

It is noted that the factors governing the development of neurological disease, host immune response, the pattern of clinical features and outcomes are poorly understood in those infected with neurotropic flaviviruses (Solomon and Winter, 2004). Therefore, studying the host response of A172 cells to WNV infection might provide some insights into the molecular mechanisms involved in mediating the neuropathogenicity observed in damaged neuronal cells. Genes which were differentially regulated were grouped into their functional groups and listed in Table 3-6.

Apoptosis is a commonly observed outcome when the virus replicates within the host cells. This active and highly conserved process of cellular self-destruction serves to

limit the spread of the virus. Apoptosis had also been shown to be a major pathway of death in mouse neuronal cells infected with another flavivirus, the dengue virus (Despres *et al.*, 1996). In this study, a group of genes causing apoptosis were found to be upregulated, and may help to shed some light on the pathways linking virus replication to apoptosis. These genes include the tumor necrosis factor superfamily (TNFSF14), nuclear factor of kappa light chain gene (NFKBIA), TNF receptor-associated factor (TRAF1) and spermidine/spermine N1-acetyltransferase (SAT). The SAT gene was recently shown to be a new mechanism in inducing apoptosis (Babbar *et al.*, 2003).

Another cause of pathological effects is due to the activation of the innate antiviral immune response pathways. Presence of double-stranded RNA (dsRNA) replication complexes from viral origins are amongst the first inducers of the immune response, and subsequently plays a role in the transcriptional activation of the IFN-α/β (or type-I IFN) pathways (Samuel, 2001). In this study on glioma cells, the activation of numerous interferon-induced proteins (such as IFIT1, IFIT2, IFI27, IFITM1, IFITM2, G1P2) lend support to this mechanism.

However, most *in vitro* studies of WNV infections were carried out on cells with no immunological functions. It was therefore difficult to examine the host cell response in native states. Glial cells are advantageous in this study as they are immune cells of CNS origins. More importantly, activated glial cells have macrophagic activity and are primed to respond to the presence of the virus, therefore allowing the display of immune-mediated neuropathology that is reflective of the native conditions in the natural CNS hosts.

In view of this, glial cells can also activate the type-II (IFN-γ) pathway and modulate the immune response by regulating cellular trafficking of various types of leukocytes, including macrophage activation and stimulation of specific T-cell cytotoxic immunity (Salmaggi *et al.,* 2002). An example of this activation was the observation that the HLA-C gene coding for the major histocompatibility complex class I (MHC-I) antigens was found to be upregulated in the A172 cells. Peptides derived from endogenous intracellular proteins are generally bound by the MHC-I molecules for presentation, thus paving the way for cell cytotoxicity in cellular immunity (Cheng *et al.*, 2004). This observation thus helped to verify the accuracy of the results. In another study on mice, the killing of target infected cells by CD8(+) T cells may explain the severe neurological pathology often observed in WNV infections (Shrestha and Diamond, 2004).

In addition, indoleamine 2,3 dioxygenase (INDO) was observed to be upregulated in the A172 cells during the infection. The INDO gene regulates catabolism of tryptophan and is also induced by IFN-γ (Du *et al.*, 2000). Increased INDO production by glial cells was found to cause neuronal injury, particularly neuroinflammatory diseases like encephalitis (Grant & Kapoor, 2003). The INDO was also found to be essential for the cytolytic activity of NK cells (Kai *et al.*, 2004), and may be another mechanism for neuronal death in WNV-infected cells. Intervention of INDO activity is a target to alleviate or prevent neuronal damage during inflammatory disease (Stone, 2002), and may be useful in treating WNV cases.

The pentaxin-related gene (PTX3) is a gene that is induced by acute inflammatory stimuli (Polentarutti *et al.*, 2000) in the brain, and this gene was also found to be upregulated in this study. This has recently been implicated in local tissue damage from various pathological factors (Bussolati *et al.*, 2003), with the attraction of inflammatory cells, resulting in an aggressive response. The PTX3 gene was also found to be rapidly produced in macrophages and was implicated in the pathogenesis of endothelial cells in atherosclerosis (Rolph *et al.*, 2002) and in causing the amplification of inflammation in innate immunity (Bottazzi *et al.*, 1997).

A major group of genes relating to the mitochondria was found to be significantly downregulated. Mitochondrial dysfunction due to respiratory-chain dysfunction and free-radical formation have been associated with neurodegenerative diseases, such as Huntington's disease, Parkinson's disease and Friedreich ataxia (Schapira, 1998). These neurological symptoms can also be observed in WNV-infected patients (Burton *et al.*, 2004) and could be the mechanism for the observed severe neurological dysfunction in the CNS.

Genes belonging to the energy synthesis pathways, consisting of the mitochondrial respiratory Kreb's cycle genes and ATP synthase genes, were decreased significantly. These included succinate dehydrogenase (SDHC), cytochrome c oxidase (COX5B and COX6B), and various genes of the ATP synthase complex (ATP5G1, ATP5C1, ATP5J, ATP5B, ATP5A1, ATP5O, and ATP5F1). The SDHC gene belongs to the complex II of the electron transport chain while COX genes codes for the terminal enzyme contributing to a proton electrochemical gradient. Mitochondrial ATP synthase catalyzes ATP synthesis by utilizing the electrochemical gradient. Decreased

or loss of energy production from the downregulation of these genes as well as overproduction of free radicals is possibly the cause of severe neurodegeneration (Eng *et al.*, 2003).

Besides being involved in respiratory functions, some mitochondrial genes are involved in cellular protection. Two of the peroxiredoxin family of antioxidant enzymes (PRDX5 and PRDX3), which help to eliminate hydrogen peroxide, were found to be down-regulated in the infected cells. This may lead to increased oxidative stress induced by reactive oxygen species (ROS), thus creating a proinflammatory condition that resulted in CNS pathology (Krapfenbauer *et al.*, 2003), and encephalitis in WNV infections. Downregulation of the PRDX3 gene was also implicated in the pathogenesis of Alzheimer's disease and Down's syndrome (Kim *et al.*, 2001).

More interestingly, downregulation of yet another gene has been associated with Alzheimer's disease and Down's syndrome. Nascent polypeptide-associated complex (NACA) helps to target nascent translated polypeptides to the ER membrane. Decreased NACA may therefore result in the mistargeting, mistranslation, and proteolysis of proteins, and might be involved in the pathology of neurodegenerative diseases (Kim *et al.*, 2002). This gene was also found to be similarly downregulated in the WNV-infected A172 cells.

Aberrations in expression of genes with putative roles in neuropathology observed in this study are listed in Table 5-1.

Besides the differentially regulated genes which are the putative cause of WNV pathology mentioned above, other genes observed in the mircoarray results has corroborated with previous findings involved in the host response during a virus infection. Examples of these include the upregulation of 2'-5'-oligoadenylate synthetase 3 (OAS3) which activates RNaseL (Presti *et al.*, 2001) in an attempt to catalyze the breakdown of the virus genome (Samuel, 2001; Lucas *et al.*, 2003). Another gene coding for the kinesin motor protein (KIF22) was also upregulated. Kinesins are microtubule-dependent molecular motors that transport organelles within cells, and are essential for the transport of WNV structural proteins to the plasma membrane (Chu & Ng, 2002a,b).

**Table 5-1.** List of differentially regulated genes involved in pathogenesis.

| ***Gene Symb*** | ***Gene Name*** | ***Fold Change*** |
|---|---|---|
| OAS3 | 2'-5'-oligoadenylate synthetase 3, 100kDa | 2.32 |
| OASL | 2'-5'-oligoadenylate synthetase-like | 3.46 |
| HLA-C | major histocompatibility complex, class I, C | 2.20 |
| INDO | indoleamine-pyrrole 2,3 dioxygenase | 3.38 |
| KIF22 | kinesin family member 22 | 2.02 |
| PTX3 | pentaxin-related gene, rapidly induced by IL-1 beta | 3.44 |
| TNFSF14 | tumor necrosis factor (ligand) superfamily, member 14 | 2.19 |
| NFKBIA | nuclear factor of kappa light polypeptide gene enhancer | 4.13 |
| TRAF1 | TNF receptor-associated factor 1 | 2.01 |
| SAT | spermidine/spermine N1-acetyltransferase | 2.18 |
| COX5B | cytochrome c oxidase subunit Vb | -2.13 |
| COX6B | cytochrome c oxidase subunit VIb | -2.41 |
| SDHC | succinate dehydrogenase complex, subunit C | -2.31 |
| PRDX5 | Peroxiredoxin 5 | -2.74 |
| PRDX3 | Peroxiredoxin 3 | -2.27 |
| ATP5G1 | ATP synthase, mitochondrial F0 complex, subunit c, isoform 1 | -2.64 |
| ATP5C1 | ATP synthase, mitochondrial F1 complex, polypeptide-γ 1 | -3.82 |
| ATP5J | ATP synthase, mitochondrial F0 complex, subunit F6 | -2.11 |
| ATP5B | ATP synthase, mitochondrial F1 complex, polypeptide-β | -2.17 |
| ATP5A1 | ATP synthase, mitochondrial F1 complex, α-subunit, isoform 1 | -2.21 |
| ATP5O | ATP synthase, mitochondrial F1 complex, O subunit | -2.00 |
| ATP5F1 | ATP synthase, mitochondrial F0 complex, subunit b, isoform 1 | -2.43 |
| NACA | nascent-polypeptide-associated complex alpha polypeptide | -2.17 |

### 5.3.2 Differences in Host Response in Different Cells May Lead to Lower Virus Yields

Besides studying the changes in gene expression between infected and mock-infected A172 cells at 24 hours p.i., a comparison was made with the less-permissive HeLa cells. The virus yield in HeLa cells was a 100-fold lower compared to A172 cells (Figure 3.5), while intracellular virus level was 4-fold lower as determined by quantitative PCR and immunofluoresnce microscopy (Tables 3-2 and 3-3). The differences in host-responses between the two types of cells were examined to determine the molecular mechanisms that might act to limit the replication of the virus in the less permissive HeLa cells. The list of genes which showed a difference in expression between the two cell types, and identified by *EASE* as significantly over-represented, are listed in Table 3-7.

Genes relating to intracellular transport functions was the predominant group found to be differentially regulated. Cytoskeletal structural constituents DNAH3 is a dynein protein that binds to microtubules to facilitate the transport of vesicles along the microtubules. Kinesin proteins (KLC2L) are also motor proteins running along microtubules, and was also found to be similarly regulated. TUBA2 and K-ALPHA-1 genes code for the tubulin subunit that polymerizes to form the microtubular complexes. Trafficking of the viral E and C proteins was previously shown to be transported along the microtubules towards the plasma membrane (Chu and Ng, 2002b), and microtubules are also involved in the trafficking of internalized virus found within endocytic vesicles (Chu and Ng, 2004). Upregulation of these genes in A172 cells over HeLa cells may therefore facilitate virus entry and production, resulting in greater virus yield in A172 cells. On the other hand, the protein tektin 3

(TEKT3) was found to be upregulated in HeLa cells. Tektin facilitates the maintenance of microtubule stability and rigidity (Roy *et al.*, 2004). The presence of this protein in higher amounts in HeLa cells may help to limit the spread of the virus within the cells, as transportation along these microtubular tracks are hindered.

Ankyrin (ANK3) is a protein that helps to link the integral membrane proteins to the underlying spectrin-actin cytoskeleton and play key roles in activities such as cell motility, activation, contact, and the maintenance of specialized membrane domains. The function of actinin (ACTN4) is to cross-link and bundle actin filaments at the plasma membrane (Honda *et al.*, 1998). Actin filaments play an essential role during the initial penetration of the virus across the plasma membrane during entry (Chu and Ng, 2004), and also for the virus maturation process during the release of virions from the plasma membrane (Chu *et al.*, 2003). Indeed, these genes were found to be upregulated in A172 cells compared to HeLa cells, thereby facilitating the virus entry and extrusion process observed. On the contrary, greater amounts of profilin were found in HeLa cells (2.5-fold more) and may therefore limit the yield of the virus. Profilin (PFN2) is involved in the regulation of actin polymerization by binding to charged actin-monomers and acts as a sequestering protein, thus modifying actin dynamics (Paavilainen *et al.*, 2004).

The West Nile virus was recently found to enter a cell through the αVβ3 integrin receptor (Chu and Ng, in press). FRMD4A is a FERM domain containing protein that mediates intermolecular interactions, including the binding and activation of the β-subunit of integrins via the talin FERM domain (Calderwood *et al.*, 2002). There was a 9-fold repression of this activating protein in HeLa cells, compared to A172 cells.

Inactivation of the integrin receptor in HeLa cells could explain the low intracellular levels of viral RNA and proteins observed, since only a small number of virus particles may be able to enter the cell successfully.

After virus attachment to the receptor, the virus is internalized by a clathrin-mediated endocytic pathway within the endosomal and lysosomal vesicles. Fusion of the enveloped virus with the lysosomal membrane occurs at low pH, resulting in the expulsion of the viral nucleocapsids into the cytoplasm around the vicinity of the ER (Chu and Ng, 2004). Syntaxin 4A (STX4A) facilitates the docking and fusion of vesicles onto membranes for the transport of intracellular materials through the t-SNARE proteins (Wyman *et al.*, 2003). ADP-ribosylation factors (ARFs) are critical in the vesicular trafficking pathway, whose activity is controlled by GTPase-activating proteins (GAPs) such as ARFGAP3. Both STX4A and ARFGAP3 function in the transport of substances between intracellular compartments mediated by transporting vesicles, and were found to be upregulated in A172 cells compared to HeLa cells. Prosaposin (PSAP) is the precursor of four sphingolipid activator proteins that are localized within lysosomes in a pH-dependent process (de Alba *et al.*, 2003), and may facilitate in virus uptake. Indeed, expression of prosaposin was found to be repressed in HeLa cells which may limit the permissiveness of the cell to the virus.

In addition to the genes involved in virus transport and packaging, genes belonging to the hexose metabolism family were also involved, particularly in glycolysis. Both aldolase (ALDOA) and glucose-6-phosphate dehydrogenase (G6PD) are housekeeping genes and thus their expression should not vary much under various physiological conditions. However, the expression of these genes was found to be

repressed in HeLa cells. On the other hand, A172 cells did not show much change in the regulation of these genes. In a study on the tick-borne encephalitis (TBE) virus, which is another flavivirus, glucose uptake was markedly increased after infection by the virulent strain rather than the attenuated strain of the virus (Starek and Babkova, 1976). The glycolytic activity was higher in cells with actively replicating viruses. A repression in glycolytic enzymes in HeLa cells observed in this study might therefore be a novel mechanism in inhibiting virus replication.

Genes belonging to the protein biosynthesis functional group is also implicated in causing the differences in permissiveness observed. The first subgroup of genes is involved in the generation of nascent peptides. Translation elongation factors (EFs) are involved in the enzymatic delivery of aminoacyl-tRNAs to the ribosome, whereas aminoacyl-tRNA synthetases catalyze the aminoacylation of tRNA by their cognate amino acid. These genes were all found to be upregulated in A172 cells compared to HeLa cells. Of the three EFs detected, translation elongation factor 1 (EEF1A2) is known to interact with the 3' stem-loop region of WNV (Blackwell and Brinton, 1997) and closely related Dengue virus (De Nova-Ocampo *et al.*, 2002) RNA to facilitate viral RNA replication. EF-1 also plays a role in actin re-organization (Ejiri, 2002; Liu *et al.*, 2002), which will affect virus entry, maturation and extrusion as mentioned above. EF-1 delta is known to decrease the metabolism of host cells to the virus's advantage by reducing the efficiency of translation of cellular mRNA, but not viral mRNA in HIV infections (Xiao *et al.*, 1998). Taken together, a decreased level of EF-1 in HeLa cells has the potential to impact the ability of WN virus to replicate efficiently inside the host. Further study on the roles of elongation factors in influencing virus replication is under study.

Ribosomal genes were also detected by *EASE* and they come under the protein biosynthesis group. Two large ribosomal proteins (RPLP2 and RPL10L) were greatly upregulated in A172 cells, while two small ribosomal proteins (RPS13 and RPS24) were upregulated in HeLa cells. This pattern of regulation is interesting although the effect of these proteins on WNV replication is not clear. However, it is believed that this is a significant finding as the capsid protein of the virus was found to localize to the nucleolus in the initial phase of infection through a novel nuclear localization signal (Wang *et al.*, 2002). The function of the nucleolus is not only for ribosome biogenesis (Shaw and Jordan, 1995), but the peri-nucleolar compartment has been implicated in RNA metabolism (Huang *et al.*, 1998). The interaction of the capsid protein with the nucleolus may therefore give rise to observed findings. It is noteworthy to mention that the localization of viral structural proteins to the nucleus can result in the induction of cytopathic effects and subsequent disease pathogenesis in the host cells (Hoyt *et al.*, 2004).

To provide further credence to the observation above, the protein importin 8 (IPO8) which mediates the nuclear import of proteins with a classical nuclear localization signal by binding to the nuclear pore complex, was found to be greatly repressed in HeLa cells. This mechanism of capsid protein entry was also observed in polyomavirus JC virus (Qu *et al.*, 2004). The capsid protein may therefore mediate the RNA processing activities within the nucleus, as it is the only viral protein found to enter the nucleus. Indeed, the surfeit-6 (SURF6) gene, which encodes a nucleolar-matrix protein with nucleic-acid binding properties (Magoulas and Fried, 2000), was found to be regulated in a similar fashion in this study. Even though the WNV is a positive-strand RNA virus and thus replicates exclusively in the cytoplasm, these

viruses are known to use the nucleus and/or nuclear components in order to facilitate their replicative processes (Hiscox, 2003). The findings here that genes involved in RNA processing within the nucleus, are also differentially regulated in cells with different permissiveness to a WNV infection, strongly supports the view that the nucleus is involved in viral replication.

Many of the genes in the RNA processing group code for protein products that are involved in nuclear RNA splicing, via the spliceosome complex in the nucleus. These include genes like SNRPB, FNBP3, SF3B5, SF4 and TRA2. As many of the genes relating to mRNA splicing mechanism were differentially regulated between the two host cells, it suggested that the host cell is adjusting its RNA processing activities not only for protein production, but also in response to the virus infection. Utilization of host splicing factors to mediate virus replication has been documented in other RNA viruses (Caputi *et al.*, 2004; Gruffat *et al.*, 2002; Momose *et al.*, 2001), but this is the first observation for a flavivirus. In the time course study of WNV infection in A172 cells, splicing factor, arginine/serine-rich 1 (SFRS1) was found to be downregulated after 18h (Appendix 10), and thus lends support to the importance of spliceosome in a flavivirus infection. SFRS1 has been found to be be pivotal in the posttranscriptional regulation of late gene expression of human papillomavirus type 16 (McPhillips *et al.*, 2004). The interaction of WNV with the nucleus is however still not well understood, and the functional significance of these interactions remains to be determined. Perturbation of the host cell nuclear functions by the capsid protein may promote an optimal environment for viral replication.

## 5.4 Progressive Global Transcriptomic Analysis of A172 Cells During WNV Infection

As previous studies and analyses were carried out at 24h post-infection, microarray studies on various time points during a WNV infection cycle were also carried out to determine the host-response changes at various stages of the WNV infection. Five different time points (i.e. 1.5h, 6h, 12h, 18h, and 24h) were used in this study, to capture the gene expression patterns at the early, middle, and late phases of infection. Various clustering methods were performed to find the best fit for the hundreds of genes being analyzed. Interesting trends in gene expression patterns during the course of WNV infection were identified, and these are presented in Section 4.4. A summary of the 6 clusters obtained is listed in Table 4-4.

Group I consists of genes whose expression were constantly upregulated throughout the course of the WNV infection. Forty-five genes were found to belong to this group, of which the immune response genes were found to be significantly over-represented by *EASE* (Appendix 7). As discussed in Section 5.3.1 above, activation of the immune response genes is the natural host response when it encounters a pathogen. It was detected as early as 1.5h after inoculation, and upregulation of these genes were maintained till the late phase of infection. The interplay between these immune-related genes is highly complex, and is made more complicated with the activation of both IFN-α/β and IFN-γ (type I and II, respectively) pathways. Some of the genes relating to the immune functions that were clustered in this group were already highlighted above, and these include the interferon-induced genes (IFITM1, IFIT1, IFIT2, and IFI27), 2',5'-oligoadenylate synthetase (OASL), and pentaxin gene (PTX3), which together may lead to the cytopathology observed in A172 cells.

Other genes of significance in this group include the N-myc (and STAT) interactor gene (NMI), bradykinin receptor B1 (BDKRB1), and protein C receptor (PROCR), all of which are involved in the inflammatory response of the host cells. Induction of these genes leads to local vasodilation, extravasation of plasma into intercellular spaces and accumulation of white blood cells and macrophages, thereby causing the observed symptom of encephalitis in patients infected with WNV.

Interestingly, genes which are antagonistic to the immune function were also found to be upregulated. Interleukin-18 binding protein (IL18BP) is activated by IFN-γ (Paulukat *et al.*, 2001) and prevents the binding of IL-18 to its receptor, thus inhibiting IL18-induced IFN-γ production (Novick *et al.*, 1999). Serine (or cysteine) proteinase inhibitor, or C1 inhibitor (SERPING1) inhibits activated C1r and C1s of the first complement component and thus regulates complement activation. These antagonistic functions serve to regulate the immune response, in view of the presence of a pro-inflammatory environment. To fully understand the immune-related effects and its relationship to the pathophysiology of infected cells, further studies are needed to be carried out in this area.

Seventeen genes were clustered into Group II, which were all constantly downregulated during the course of the WNV infection. Even though the functions of some genes are listed in Appendix 8, no specific ontology group was found to be significantly over-represented by *EASE*. Therefore, even though 3 genes relating to cell proliferation were found to be downregulated, it does not necessarily signify that the process of cell multiplication is totally shut down. The same is true for all 5

categories listed. Nevertheless, some of these constantly downregulated genes are of significance in the virus replication process.

Coatomer protein complex, zeta (COPZ1), is a protein involved in intracellular transport, and is closely associated with the Golgi apparatus and the clathrin vesicle coat. Downregulation of this gene constantly is interesting, as the virus is known to enter the cell through a clathrin-mediated endocytic pathway (Chu and Ng, 2004). Further studies in the genes related to the ER and Golgi apparatus might help to further elucidate the relationship of these organelles with virus replication. Two genes relating to carbohydrate metabolism, alpha glucosidase II alpha subunit (GANAB) and succinate dehydrogenase complex (SDHC), were also found to be down-regulated. The significance of a downregulation in carbohydrate metabolism and respiration has been discussed in Section 5.3.

Genes that were upregulated at the early phase of infection (i.e. at 1.5h p.i.), but were subsequently downregulated after the mid-phase of infection (6 h p.i.) were found in Group III (Figure 4-21). There were 23 genes in this group, and the dominant gene ontology group belongs to cell cycle regulation. Of the 4 genes relating to the cell cycle regulation, NHP2 non-histone chromosome protein 2-like 1 (NHP2L1) is also part of the ribosome, and is intricately linked to the nucleolus and spliceosome complex (Leung and Lamond, 2002). Cell cycle genes appeared in a significant number of instances throughout this study, but the patterns of expression has been sporadic. The relationship of these genes on WNV infection is not clear, but it is noted that changes in cell cycle regulation has been reported in other virus systems

such as cytomegalovirus (Hertel and Morcarski, 2004) and HIV (Mannioui *et al.*, 2004).

Genes in Group IV were initially upregulated, but were found to be downregulated after 18 h post-infection during the late phase of infection (Figure 4-22). Eight genes belonging to nucleic acid metabolism were significantly over-represented as determined by *EASE*, with six of these genes also belonging to the DNA-dependent transcription regulation ontology group. It was noted that after 18 hours post-infection, the WNV entered the late phase of infection with peak virus production (Ng *et al.*, 2001). While the host-cells were still viable at this stage of the virus replication process, it was observed that various cellular genes can be downregulated through regulating the activity of the transcription factors. Of these genes, YY1 transcription factor (YY1) is of interest as accumulation of this protein within the nucleus during the course of a virus infection was found to decrease virus yield in vaccinia virus (Slezak *et al.*, 2004).

Also of interest is the presence of 2 kinesin (KIF3C and RAB6KIFL) and nexilin (F-acting binding; NEXN) genes which were also found to be upregulated during the early/mid phases of infection, but were downregulated after 18h. As discussed in Section 5.3.2, kinesins are an important group of motor proteins involved in the intracellular transport of viral proteins. Upregulation of these genes initially will therefore facilitate in viral proteins transport, but downregulation of the motor proteins at the late phase of infection may denote that most of the viral particles have been assembled, or that sites of replication have been established.

Genes in Group V show the exact opposite in expression pattern from Group IV. They were downregulated during the initial phases of infection, but were upregulated after 18h post-infection (Figure 4-23). There are 27 genes in this group, and 2 ontology groups belonging to cellular physiological processes and integral membrane proteins were found to be significantly over-represented by *EASE*.

Three genes from the cellular physiological processes group were found to be of interest here. Death-associated protein kinase 3 (DAPK3) and pleckstrin homology-like domain (PHLDA2) genes are both involved in apoptosis, and both were found to be upregulated after 18 h p.i. DAPK3 induces morphological changes in apoptosis when it is overexpressed in mammalian cells through a nuclear pathway (Kawai *et al.*, 2003), while PHLDA2-induced apoptosis is Fas-mediated (Qian *et al.*, 1997). The role of apoptosis in WNV infection and other genes involved in cell death were also discussed in Section 5.3.1. Laminin alpha 1 (LAMA1) is involved in basement membrane formation through the formation of glycoproteins and helps in the regulation of cell adhesion. This gene was found to be upregulated after 18h p.i., and was also similarly found to be heavily upregulated in cells afflicted with Alzheimer's disease and playing a role in its pathophysiology (Palu and Liesi, 2002).

In the integral membrane proteins groups, interleukin 27 receptor (WSX1) was also upregulated after 18h p.i. WSX1 binds to IL-27 and is important for the activation of STAT1 in the immune response (Takeda *et al.*, 2003). In yet another link to Alzheimer's Disease, low density lipoprotein-related protein (LRP1) which is involved in the transport of amyloid-β, was also found to be upregulated after 18h p.i. (Deane *et al.*, 2004).

The gene expression pattern in Group VI is rather unique when compared to the rest of the groups. The genes were all downregulated during the WNV infection, except at 6 h p.i. whereby they were upregulated (Figure 4-24). Interestingly, this group has the most number of genes (i.e. 51 genes) and 3 functional groups were found to be significantly over-represented by *EASE*.

Genes belonging to the endoplasmic reticulum group were highly represented. Reticulon 4 (RTN4), or NOGO protein, is involved in membrane trafficking. Upregulation of this gene in neurons has been implicated in epilepsy (Bandtlow *et al.*, 2004) and in inducing apoptosis (Qi *et al.*, 2003). Emopamil-binding protein (EBP) is an integral membrane protein of the endoplasmic reticulum involved in the trafficking of drugs (Hanner *et al.*, 1995) and possesses sterol-isomerase activity (Braverman *et al.*, 1999). Another gene involved in sterol metabolism is the sterol regulatory element binding transcription factor 1 (SREBF1). The nascent transcription factor is localized at the ER before entering the nucleus to bind at the sterol regulatory element-1 (SRE1) (Wang *et al.*, 1994). Kinesin-associated protein 3 (KIFAP3) is concentrated around the ER, and can serve as an adapter to kinesin II in the transport of vesicles along microtubules (Shimizu *et al.*, 1996). The SEC23 interacting protein (SEC23IP) has been shown to cause disorganization of the endoplasmic reticulum-Golgi intermediate compartment and Golgi apparatus, which suggests its role in the early secretory pathway (Tani *et al.*, 1999).

The upregulation of these ER related genes at 6h p.i. is highly significant as the trafficking of the uncoated viral RNA genomes to the ER and subsequent virus maturation occurs at the mid-early phase of infection (Chu and Ng, 2004). The genes

are therefore important for virus uptake and maturation. Another gene that is involved in intracellular protein transport is the sorting nexin 3 (SNX3) gene through the endocytic process (Haft *et al.*, 1998).

Genes involved in the cytoskeletal structures include tropomyosin 3 and tropomyosin 4 (TPM3 and TPM4), which are molecular motors that bind to actin filaments. Interestingly, actin dynamics may be mediated by the abl-interactor 1 (ABI1) gene, which was also upregulated at 6h p.i. ABI1 increases actin polymerization activity and is involved in Rac-dependent lamellipodia protrusion (Innocenti *et al.*, 2004). Such lamellipodia protrusions were also observed in the WNV-infected cells (Lee and Ng, 2004).

Genes relating to the mitochondria were also again highly represented. Defects in mitochondrial function have implications in neuropathology, and this was discussed in Section 5.3.1. However, the mitochondria genes listed in Group VI were different from those above, albeit with similar functions. Succinate-CoA ligase (SUCLA2), isocitrate dehydrogenase 1 (IDH1), cytochrome c oxidase (COX6C), surfeit 1 (SURF1) and ATP synthase (ATP5F1) are all involved in the respiratory chain. These genes were upregulated at 6h p.i., but were again downregulated at the late phase of infection. The change in metabolic activity at this time point of infection is therefore be interesting. Abberations in SURF1 and COX genes have been shown to cause Leigh syndrome and brain damage (Moslemi, *et al.*, 2003). Sterol carrier protein 2 (SCP2) is involved in the transfer of sterols between plasma membranes, lysosomes, peroxisomes, mitochondria, and the endoplasmic reticulum, and also helps in the protection against oxidative damage (Dansen *et al.*, 2004).

## 5.5 Conclusion

Genes relating to the apoptosis and the immune response were constantly highlighted in this study. This was naturally so given the nature of the assault suffered by the host cells during a virus infection. However, there were other functional classes which were also highly significant in this study. These included the interplay between the cytoskeletal structures and intracellular transport genes in mediating virus replication and maturation. Genes involved in the nucleolar, spliceosome, ribosomal and protein synthesis functional groups were also highly significant and warrant further analysis. Defects in mitochondrial function and its effects on cellular metabolism were also significant. Finally, many of these genes were observed to be related to the pathophysiology of other neuronal diseases like Parkinson's and Alzheimer's diseases, and may represent a common pathway in the degeneration of the cells in the CNS.

The results of this extensive study on the global analyses of WNV-host interaction have been interesting, and displayed a complex network of host responses that may affect the virus replication during the infection process. This has helped to provide new insights into the pathogenesis of the diseases. The use of glial A172 cells, which are brain cells with immune functions, has shed some important insights into the molecular mechanisms behind the observed neuronal pathology. The use of the less permissive HeLa cells has also helped in our understanding on which host factors are important in limiting virus replication.

# REFERENCES

**Amersham Biosciences. 2002.** *Microarray Handbook*, Amersham Biosciences Corp, NJ.

**Andrews DM, Matthews VB, Sammels LM. 1999.** The severity of Murray Valley encephalitis in mice is linked to neutrophil infiltration and inducible nitric oxide synthetase activity in the central nervous system. *J Virol.* **73**:8781-8790.

**Arturo LA, Fratkin J, Stokic DS, Harrington T, Webb RM, Slavinsky SA. 2003.** West Nile poliomyelitis. *Lancet Infect Dis*. **3**:9–10.

**Babbar N, Ignatenko NA, Casero RA, Gerner EW. 2003.** Cyclooxygenase-independent induction of apoptosis by sulindac sulfone is mediated by polyamines in colon cancer. *J Biol Chem.* **278**:47762-47775.

**Bandtlow CE, Dlaska M, Pirker S, Czech T, Baumgartner C, Sperk G. 2004.** Increased expression of Nogo-A in hippocampal neurons of patients with temporal lobe epilepsy. *Eur J Neurosci.* **20**:195-206.

**Beasley DW, Davis CT, Whiteman M, Granwehr B, Kinney RM, Barrett AD. 2004.** Molecular determinants of virulence of West Nile virus in North America. *Arch Virol Suppl.* **18**:35-41.

**Belot N, Pochet R, Heizmann CW, Kiss R, Decaestecker C. 2002.** Extracellular S100A4 stimulates the migration rate of astrocytic tumor cells by modifying the organization of their actin cytoskeleton. *Biochim Biophys Acta*. **1600**:74-83.

**Berthet FX, Zeller HG, Drouet MT, Rauzier J, Digoutte JP, Deubel V. 1997.** Extensive nucleotide changes and deletions within the envelope glycoprotein gene of Euro-African West Nile viruses. *J Gen Virol.* **78**:2293-2297.

**Bernard KA, Maffei JG, Jones SA, Kauffman EB, Ebel GD. 2001.** West Nile virus

infection in birds and mosquitoes, New York State, 2000. *Emerg Infect Dis.* **7**:679-685.

**Bilban M, Buehler LK, Head S, Desoye G, Quaranta V. 2002.** Normalizing DNA microarray data. *Curr Issues Mol Biol.* **4**:57-64.

**BioRad. 2004.** Standard curves as troubleshooting tools for real-time RT-PCR assays. *BioRadiations* **112**:21-22

**Blackwell JL, Brinton MA. 1997.** Translation elongation factor-1 alpha interacts with the 3' stem-loop region of West Nile virus genomic RNA. *J Virol.* **71**:6433-6444.

**Bottazzi B, Vouret-Craviari V, Bastone A, De Gioia L, Matteucci C, Peri G, Spreafico F, Pausa M, D'Ettorre C, Gianazza E, Tagliabue A, Salmona M, Tedesco F, Introna M, Mantovani A. 1997.** Multimer formation and ligand recognition by the long pentraxin PTX3: similarities and differences with the short pentraxins C-reactive protein and serum amyloid P component. *J Biol Chem.* **272**: 32817–32823

**Braverman N, Lin P, Moebius FF, Obie C, Moser A, Glossmann H, Wilcox WR, Rimoin DL, Smith M, Kratz L, Kelley RI, Valle D. 1999.** Mutations in the gene encoding 3 beta-hydroxysteroid-delta 8, delta 7-isomerase cause X-linked dominant Conradi-Hanermann syndrome. *Nat Genet.* **22**:291-294.

**Brazma A, Vilo J. 2000.** Gene expression data analysis. *FEBS Lett.* **480**:17-24.

**Briese T, Rambaut A, Pathmajeyan M, Bishara J, Weinberger M, Pitlik S, Lipkin WI. 2002.** Phylogenetic analysis of a human isolate from the 2000 Israel West Nile virus epidemic. *Emerg Infect Dis.* **8**:528-531.

**Brinton MA. 1986**. Replication of flaviviruses. In: Schlesinger S, Schlesinger M, editors. *Togaviridae and Flaviridae* The Viruses. Plenum, New York. p. 329-376.

**Brinton MA. 2002.** The molecular biology of West Nile Virus: A new invader of the

western hemisphere. *Annu Rev Microbiol.* **56**:371-402.

**Brinton MS, Kurane I, Mathew A, Zeng L, Shi PY, Rothman A, Ennis FA. 1998.** Immune medicated and inherited defences against flaviviruses. *Clin Diagn Virol.* **10**:129-139.

**Burke DS, Morill JC. 1987.** Levels of interferon in the plasma and cerebrospinal fluid of patients with acute Japanese encephalitis. *J Infect Dis* **155**:797-799.

**Burke DS, Monath TP. 2001**. Flaviviruses. In *Fields Virology,* ed. Knipe DM, Howley PM. Philadelphia: Lippin-cott, Williams & Wilkins. 4th ed. , pp. 1043-125

**Burton JM, Kern RZ, Halliday W, Mikulis D, Brunton J, Fearon M, Pepperell C, Jaigobin C. 2004.** Neurological manifestations of West Nile virus infection. *Can J Neurol Sci.* **31**:185-193.

**Bussolati B, Peri G, Salvidio G, Verzola D, Mantovani A, Camussi G. 2003.** The long pentraxin PTX3 is synthesized in IgA glomerulonephritis and activates mesangial cells. *J Immunol.* **170**:1466-1472.

**Butt E, Gambaryan S, Gottfert N, Galler A, Marcus K, Meyer HE. 2003**. Actin binding of human LIM and SH3 protein is regulated by cGMP- and cAMP-dependent protein kinase phosphorylation on serine 146. *J Biol Chem*. **278**:15601-15607.

**Calderwood DA, Yan B, de Pereda JM, Alvarez BG, Fujioka Y, Liddington RC, Ginsberg MH. 2002.** The phosphotyrosine binding-like domain of talin activates integrins. *J Biol Chem.* **277**:21749-21758.

**Calisher CH, Karabatsos N, Dalrymple JM, Shope RE, Porterfield JS. 1989.** Antigenic relationships between flaviviruses as determined by cross-neutralization tests with polyclonal antisera. *J Gen Virol.* **70**:37-43.

**Cantile C, Del Piero F, Di Guardo G, Arispici M. 2001.** Pathologic and immunohistochemical findings in naturally occurring West Nile virus infection in

horses. *Vet Pathol.* **38**:414–421.

**Caputi M, Freund M, Kammler S, Asang C, Schaal H. 2004.** A bidirectional SF2/ASF- and SRp40-dependent splicing enhancer regulates human immunodeficiency virus type 1 rev, env, vpu, and nef gene expression. *J Virol.* **78**:6517-6526.

**Catteau A, Kalinina O, Wagner MC, Deubel V, Courageot MP, Despres P. 2003** Dengue virus M protein contains a proapoptotic sequence referred to as ApoptoM. *J Gen Virol.* **84**:2781–2793.

**Ceccaldi PE, Lucas M, Despres P. 2004.** New insights on the neuropathogenicity of West Nile virus. *FEMS Microbiol Lett.* **233**:1-6.

**Celis JE, Kruhoffer M, Gromova I, Frederiksen C, Ostergaard M, Thykjaer T, Gromov P, Yu J, Palsdottir H, Magnusson N, Orntoft TF. 2000.** Gene expression profiling: monitoring transcription and translation products using DNA microarrays and proteomics. *FEBS Lett.* **480**:2-16.

**Centers for Disease Control and Prevention (CDC). 2004**. 2003 West Nile virus activity in the United States (reported as of May 21, 2004). www.cdc.gov.

**Chang YE, Laimin LA. 2000.** Microarray analysis identifies interferon-inducible genes and Stat-1 as major transcriptional targets of human papillomavirus type 31. *J Virol.* **74**:4174–4182.

**Cheng Y, King NJ, Kesson AM. 2004.** Major histocompatibility complex class I (MHC-I) induction by West Nile virus: involvement of 2 signaling pathways in MHC-I up-regulation. *J Infect Dis.* **189**:658-668.

**Chowers MY, Lang R, Nasser F, David D, Giladi M. 2001.** Clinical characteristics of the West Nile fever outbreak, Israel 2000. *Emerg Infect Dis.* **7**:611-614.

**Chu JJH, Ng ML. 2002a.** Infection of polarized epithelial cells with flavivirus West

Nile: polarized entry and egress of virus occur through the apical surface. *J Gen Virol.* **83**:2427-2435.

**Chu JJH, Ng ML. 2002b.** Trafficking mechanism of West Nile (Sarafend) virus structural proteins. *J Med Virol.* **67**:127-136.

**Chu JJH, Ng ML. 2003a**. Characterization of a 105-kDa plasma membrane associated glycoprotein that is involved in West Nile virus binding and infection. *Virology.* **312**:458-469.

**Chu JJH, Ng ML. 2003b.** The mechanism of cell death during West Nile virus infection is dependent on initial infectious dose. *J Gen Virol.* **84**:3305-3314.

**Chu JJ, Choo BG, Lee JW, Ng ML. 2003.** Actin filaments participate in West Nile (Sarafend) virus maturation process. *J Med Virol.* **71**:463-472.

**Chu JJ, Ng ML. 2004.** Infectious entry of West Nile virus occurs through a clathrin-mediated endocytic pathway. *J Virol.* **78**:10543-10555.

**Chu JJ, Ng ML. (In press.)** Interaction of West Nile virus with aVβ3 integrin mediates virus entry into cells. *J Biol Chem.*

**Chuaqui RF, Bonner RF, Best CJM, Gillespie JW, Flaig MJ, Hewitt SM, Phillips JL, Krizman DB, Tangrea MA, Ahram M, Linehan WM, Knezevic V, Emmert-Buck MR. 2002.** Post-analysis follow-up and validation of microarray experiments. *Nature Genet Supp.* **32**:509-514.

**Churchill, G.A. 2002.** Fundamentals of experimental design for cDNA microarrays. *Nature Genet.* **32**:490–495.

**Corbeil J, Sheeter D, Rought S, Du P, Ferguson M, Masys DR. 1999.** Magnitude and specificity of temporal gene expression during HIV-1 infection of a CD4+ T cell. *Nature Genet.* **23**:39-40.

**Cummings CA, Relman DA. 2000**. Using DNA Microarrays to study host-microbe

interactions. *Emerging Infect Dis.* **6**:513-525

**Dansen TB, Kops GJ, Denis S, Jelluma N, Wanders RJ, Bos JL, Burgering BM, Wirtz KW. 2004.** Regulation of sterol carrier protein gene expression by the forkhead transcription factor FOXO3a. *J Lipid Res.* **45**:81-88.

**De Alba E, Weiler S, Tjandra N. 2003.** Solution structure of human saposin C: pH-dependent interaction with phospholipid vesicles. *Biochemistry.* **42**:14729-14740.

**De Nova-Ocampo M, Villegas-Sepulveda N, del Angel RM. 2002.** Translation elongation factor-1alpha, La, and PTB interact with the 3' untranslated region of dengue 4 virus RNA. *Virology.* **295**:337-347.

**Deane R, Wu Z, Sagare A, Davis J, Du Yan S, Hamm K, Xu F, Parisi M, LaRue B, Hu HW, Spijkers P, Guo H, Song X, Lenting PJ, Van Nostrand WE, Zlokovic BV. 2004.** LRP/amyloid beta-peptide interaction mediates differential brain efflux of Abeta isoforms. *Neuron.* **43**:333-344.

**Despres P, Fiamand M, Ceccaldi PE. 1996**. Human isolates of dengue type 1 induce apoptosis in mouse neuroblastoma cells. *J Virol.* **70**:4090-4096.

**Detweiler CS, Cunanan DB, Falkow S. 2001.** Host microarray analysis reveals a role for the Salmonella response regulator phoP in human macrophage cell death. *Proc Natl Acad Sci. U S A.* **98**:5850-5855.

**Diamond MS, Shrestha B, Marri A, Mahan D, Engle M. 2003a** B cells and antibody play critical roles in the immediate defense of disseminated infection by West Nile virus encephalitis. *J Virol*. **77**:2578–2586.

**Diamond MS, Sitati EM, Friend LD, Higgs S, Shrestha B, Engle M. 2003b.** A critical role for induced IgM in the protection against West Nile virus infection. *J Exp Med.* **198**:1853–1862.

**Douglas MW, Kesson AM, King NJC. 1994.** CTL recognition of West Nile virus-

infected fibroblasts is cell cycle dependent and is associated with virus-induced increases in class I MHC antigen expression. *Immunology.* **82**:561-570.

**Du MX, Sotero-Esteva WD, Taylor MW. 2000.** Analysis of transcription factors regulating induction of indoleamine 2,3-dioxygenase by IFN-gamma. *J Interferon Cytokine Res.* **20**:133-142.

**Dunster LM, Gibson CA, Stephenson JR, Minor PD, Barrett AD. 1990.** Attenuation of virulence of flavivirus following passage in HeLa cells. *J Gen Virol.* **71**:601-607.

**Eisen MB, Spellman PT, Brown PO, Botstein D. 1998.** Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A.* **95**:14863-14868.

**Ejiri S. 2002.** Moonlighting functions of polypeptide elongation factor 1: from actin bundling to zinc finger protein R1-associated nuclear localization. *Biosci Biotechnol Biochem.* **66**:1-21.

**Eng C, Kiuru M, Fernandez MJ, Aaltonen LA. 2003**. A role for mitochondrial enzymes in inherited neoplasia and beyond. *Nat Rev Cancer.* **3**:193-202.

**Engle MJ, Diamond MS. 2003.** Antibody prophylaxis and therapy against West Nile virus infection in wild-type and immunodeficient mice. *J Virol*. **77**:12941–12949.

**Fredericksen BL, Smith M, Katze MG, Shi PY, Gale M Jr. 2004.** The host response to West Nile Virus infection limits viral spread through the activation of the interferon regulatory factor 3 pathway. *J Virol.* **78**:7737-7747.

**Geiss GK, Bumgarner RE, An MC, Agy MB, van't Wout AB, Hammersmark E. 2000.** Large-scale monitoring of host cell gene expression during HIV-1 infection using cDNA microarrays. *Virology.* **266**:8-16.

**Geiss GK, An MC, Bumgarner RE, Hammersmark E, Cunningham D, Katze MG. 2001a.** Global impact of influenza virus on cellular pathways is mediated by

both replication-dependent and -independent events. *J Virol.* **75**:4321–4331.

**Geiss GK, Bumgarner RE, An MC, Agy MB, van 't Wout AB, Hammersmark E, Carter VS, Upchurch D, Mullins JI, Katze MG. 2001b.** Large-scale monitoring of host cell gene expression during HIV-1 infection using cDNA microarrays. *Virology.* **266**:8–16.

**George S, Gourie-Devi M, Rao JA, Prasad SR, Pavri KM. 1984.** Isolation of West Nile virus from brains of children who have died of encephalitis. *Bull World Health Organ.* **62**:879-882.

**Giard DJ, Aaronson SA, Todaro GJ, Arnstein P, Kersey JH, Dosik H, Parks WP. 1973.** *In vitro* cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. *J Natl Cancer Inst.* **51**:1417-1423.

**Goodin JL, Rutherford CL. 2002.** Characterization of human ribosomal S3a gene expression during adenosine 3':5' cyclic monophosphate induced neuroendocrine differentiation of LNCaP cells. Regulation of S3a gene expression in LNCaP. *Mol. Biol. Rep*. **29**:301-316.

**Grant GR, Manduchi E, Pizarro A, Stoeckert CJ. 2003.** Maintaining data integrity in microarray data management. *Biotech Bioeng.* **84**:795-800.

**Grant R, Kapoor V. 2003.** Inhibition of indoleamine 2,3-dioxygenase activity in IFN-gamma stimulated astroglioma cells decreases intracellular NAD levels. *Biochem Pharmacol.* **66**:1033-1036.

**Gruffat H, Batisse J, Pich D, Neuhierl B, Manet E, Hammerschmidt W, Sergeant A. 2002.** Epstein-Barr virus mRNA export factor EB2 is essential for production of infectious virus. *J Virol.* **76**:9635-9644.

**Haft CR, de la Luz Sierra M, Barr VA, Haft DH, Taylor SI. 1998.** Identification of a family of sorting nexin molecules and characterization of their association with

receptors. *Mol Cell Biol.* **18**:7278-7287.

**Hanner M, Moebius FF, Weber F, Grabner M, Striessnig J, Glossmann H. 1995.** Phenylalkylamine Ca2+ antagonist binding protein. Molecular cloning, tissue distribution, and heterologous expression. *J Biol Chem.* **270**:7551-7557.

**Hase T, Summers PL, Eckels KH, Baze WB. 1989.** Maturation process of Japanese Encephalitis virus in cultured mosquito cells *in vitro* and mouse brain cells *in vivo*. *Arch Virol.* **96**:135-151.

**Hayes EB, O'Leary DR. 2004.** West Nile virus infection: a pediatric perspective. *Pediatrics.* **113**:1375-1381.

**Heinz FX, Roehrig JT. 1990.** Flaviviruses in *Immuno-chemistry of Viruses II. The basis for Serodiagnosis and Vaccines.* MHV can Regenmortel and AR Neurath, Eds. Elsevier, Amsterdam. pp 289-305.

**Heinz FX, Stiasny K, Puschner-Auer G, Holzmann H, Allison SL. 1994.** Structural changes and functional control of the tick-borne encephalitis virus glycoprotein E by the heterodimeric association with protein prM. *Virology.* **198**:109-117.

**Heinz FX, Allison SL. 2000.** Structures and mechanisms in flavivirus fusion. *Adv Virus Res.* **55**:231-269.

**Heinz FX, Collett MS, Purcell RH, Gould EA, Howard CR, Houghton M, Moormann RJM, Rice CM, Thiel HJ. 2000.** *Flaviviridae*. In: van Regenmortel, M.H.V., Fauquet, C.M., Houghton, M., Bishop, D.H.L., Carstens, E.B., Estes, M.K., Lemon, S.M., Maniloff, J., Mayo, M.A., McGeoch, D.J., Pringle, C.R., Wickner, R.B. (eds) Virus Taxonomy. Classification and nomenclature of viruses. 7th Report of the International Committee for the Taxonomy of Viruses. Academic Press, San Diego, pp. 859-878.

**Herrero J, Valencia A, Dopazo J. 2001.** A hierarchical unsupervised growing neural network for clustering gene expression patterns. *Bioinformatics.* **17**:126-136.

**Hertel L, Mocarski ES. 2004.** Global analysis of host cell gene expression late during cytomegalovirus infection reveals extensive dysregulation of cell cycle gene expression and induction of Pseudomitosis independent of US28 function. *J Virol.* **78**:11988-12011.

**Hess KR, Zhang W, Baggerly KA, Stivers DN, Coombes KR. 2001.** Microarrays: handling the deluge of data and extracting reliable information. *Trends Biotech.* **19**:463-468.

**Hiscox JA. 2003.** The interaction of animal cytoplasmic RNA viruses with the nucleus to facilitate replication. *Virus Res*. **95**:13-22.

**Hoffman R, Seidl T, Dugas M. 2002.** Profound effect of normalization on detection of differentially expressed genes in oligonucleotide microarray data analysis. *Genome Biol.* **3**:1-11.

**Holloway AJ, van Laar RK, Tothill RW, Bowtell DDL. 2002.** Options available – from start to finish – for obtaining data from DNA microarrays II. *Nature Genet supp.* **32**:481-489.

**Honda, K., Yamada, T., Endo, R., Ino, Y., Gotoh, M., *et al.*, 1998.** Actinin-4, a novel actin-bundling protein associated with cell motility and cancer invasion. *J Cell Biol.* **140**:1383-1393.

**Hosack DA, Dennis G, Sherman BT, Lane H, Lempicki RA. 2003.** Identifying biological themes within lists of genes with *EASE*. *Genome Biol.* **4:**R70.

**Hoyt CC, Bouchard RJ, Tyler KL. 2004.** Novel nuclear herniations induced by nuclear localization of a viral protein. *J Virol.* **78**:6360-6369.

**Huang S, Deerinck TJ, Ellisman MH, Spector L. 1998.** The perinucleolar

compartment and transcription. *J Cell Biol.* **143**:35-47.

**Huang SH, Triche T, Jong AY. 2002**. Infectomics: genomics and proteomics of microbial infections. *Funct Integr Genomics.* **1**:331-344.

**Ilkal MA, Mavale MS, Prasanna Y, Jacob PG, Geevarghese G. 1997.** Experiment studies on the vector potential of certain Culex species to West Nile Virus. *Indian J Med Res.* **106**:225-228.

**Innocenti M, Zucconi A, Disanza A, Frittoli E, Areces LB, Steffen A, Stradal TE, Di Fiore PP, Carlier MF, Scita G. 2004**. Abi1 is essential for the formation and activation of a WAVE2 signalling complex. *Nat Cell Biol.* **6**:319-327.

**Jia XY, Briese T, Jordan I, Rambaut A, Chi HC, Mackenzie JS, Hall RA, Scherret J, Lipkin WI. 1999.** Genetic analysis of West Nile New York 1999 encephalitis virus. *Lancet.* **354**:1971-1972.

**Johnson CM, Yang S, Sellins KS, Frank GR. 2004.** Selection of HPRT primers as controls for determination of mRNA expression in dogs by RT-PCR. *Vet Immunol Immunopathol.* **99**:47-51.

**Johnston C, Jiang W, Chu T, Levine B. 2001.** Identification of genes involved in the host response to neurovirulent alphavirus infection. *J Virol.* **75**:10431–10445.

**Kai S, Goto S, Tahara K, Sasaki A, Tone S, Kitano S. 2004.** Indoleamine 2,3-dioxygenase is necessary for cytolytic activity of natural killer cells. *Scand J Immunol.* **59**:177-182.

**Kawai T, Akira S, Reed JC. 2003.** ZIP kinase triggers apoptosis from nuclear PML oncogenic domains. *Mol Cell Biol.* **23**:6174-6186.

**Kesson, AM, King NJ. 2001.** Transcriptional regulation of major histocompatibility complex class I by flavivirus West Nile is dependent on NF-kappa B activiation. *J Infect Dis.* **15**:947-954.

**Kim SH, Fountoulakis M, Cairns N, Lubec G. 2001.** Protein levels of human peroxiredoxin subtypes in brains of patients with Alzheimer's disease and Down syndrome. *J Neural Transm Suppl.* **61**:223-235.

**Kim SH, Shim KS, Lubec G. 2002.** Human brain nascent polypeptide-associated complex alpha subunit is decreased in patients with Alzheimer's disease and Down syndrome. *J Investig Med.* **50**:293-301

**King NJC, Mullbacher A, Tian L, Rodger J, Lidbury B, Tha Hla R. 1993.** WNV infection induces susceptibility of *in vitro* outgrowth murine blastocysts to specific lysis by parentally directed alloimmune and virus-immune cytotoxic T cells. *J Reprod Immunol.* **23**:131-144.

**Knudsen S. 2002.** A biologist's guide to analysis of DNA microarray data. John Wiley & Sons, New York, NY.

**Ko KK, Igarashi A, Fukai K. 1979.** Electron microscopic observations on *Aedes albopictus* cells infected with dengue viruses. *Arch Virol.* **62**:41-52.

**Krapfenbauer K, Engidawork E, Cairns N, Fountoulakis M, Lubec G. 2003.** Aberrant expression of peroxiredoxin subtypes in neurodegenerative disorders. *Brain Res.* **967**:152-160.

**Kuhn RJ, Zhang W, Rossmann MG, Pletnev SV, Corver J, Lenches E, Jones CT, Mukhopadhyay S, Chipman PR, Strauss EG, Baker TS, Strauss JH. 2002.** Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. *Cell.* **108**:717–725

**Lanciotti RS, Ebel GD, Deubel V, Kerst AJ, Murri S, Meyer R, Bowen M, McKinney N, Morrill WE, Crabtree MB, Kramer LD, Roehrig JT. 2002.** Complete genome sequences and phylogenetic analysis of West Nile virus strains isolated from the United States, Europe, and the Middle East. *Virology.* **298**:96-105.

**Lee JW, Ng ML. 2004.** A nano-view of West Nile virus-induced cellular changes during infection. *J Nanobiotechnology.* 2:6.

**Leung AK, Lamond AI. 2002**. In vivo analysis of NHPX reveals a novel nucleolar localization pathway involving a transient accumulation in splicing speckles. *J Cell Biol.* **157**:615-629.

**Li J, Loeb JA, Shy ME, Aashit KS, Tselis AC, Kupski WJ, Lewis RA. 2003.** Asymmetric flaccid paralysis: a neuromuscular presentation of West Nile virus infection. *Ann Neurol.* **53**:703–710.

**Lindenbach BD, Rice CM. 1999.** Genetic interaction of flavivirus nonstructural proteins NS1 and NS4A as a determinant of replicase function. *J Virol.* **73**:4611-4621.

**Liu G, Grant WM, Persky D, Latham VM Jr, Singer RH, Condeelis J. 2002.** Interactions of elongation factor 1alpha with F-actin and beta-actin mRNA: implications for anchoring mRNA in cell protrusions. *Mol Biol Cell.* **13**:579-592.

**Livak KJ, Schmittgen TD. 2001.** Analysis of relative gene expression using real-time quantitative PCR and the $2^{-\Delta\Delta C}{}_T$ method. *Methods* **25**:402-408.

**Lucas M, Mashimo T, Frenkiel MP, Simon-Chazottes D, Montagutelli X, Ceccaldi PE, Guenet JL, Despres P. 2003.** Infection of mouse neurones by West Nile virus is modulated by the interferon-inducible 2'-5' oligoadenylate synthetase 1b protein. *Immunol Cell Biol.* **81**:230-236.

**Mackenzie JS, Barrett AD, Deubel V. 2002.** The Japanese encephalitis serological group of flaviviruses: a brief introduction to the group. *Curr Top Microbiol Immunol.* **267**:1-10.

**Magoulas C, Fried M. 2000.** Isolation and genomic analysis of the human surf-6 gene: a member of the Surfeit locus. *Gene.* **243**:115-123.

**Manger ID, Relman DM. 2000.** How the host 'sees' pathogens: global gene

expression responses to infection. *Curr Opin Immunol.* **12**:215-218.

**Mannioui A, Schiffer C, Felix N, Nelson E, Brussel A, Sonigo P, Gluckman JC, Canque B. 2004.** Cell cycle regulation of human immunodeficiency virus type 1 integration in T cells: antagonistic effects of nuclear envelope breakdown and chromatin condensation. *Virology.* **329**:77-88.

**Marsh M, McMahon HT. 1999.** The structural era of endocytosis. *Science*. **85:**215-220.

**Martin DA, Biggerstaff BJ, Allen B, Johnson AJ, Lanciotti RS, Roehrig JT. 2002.** Use of immunoglobuilin M cross-reactions in differential diagnosis of human flaviviral encephalitis infections in the United States. *Clin Diagn Lab Immunol.* **9**:544-549.

**Mason PW. 1989.** Maturation of Japanese encephalitis virus glycoproteins produced by infected mammalian and mosquito cells. *Virology.* **169**:354-364.

**Matsumura T, Shiraki K, Sashikata T, Hotta S. 1977.** Morphogenesis of dengue-1 virus in cultures of a human leukemic leukocyte line (J-111). *Microbiol Immunol.* **21**:329-334.

**McPhillips MG, Veerapraditsin T, Cumming SA, Karali D, Milligan SG, Boner W, Morgan IM, Graham SV. 2004.** SF2/ASF binds the human papillomavirus type 16 late RNA control element and is regulated during differentiation of virus-infected epithelial cells. *J Virol.* **78**:10598-10605.

**Miller BR, Nasci RS, Godsey MS, Savage HM, Lutwama JJ. 2000.** First field evidence for natural vertical transmission of West Nile virus in Culex uni-vittatus complex mosquitoes from Rift Valley province. *Am J Trop Med Hyg.* **62**:240-246.

**Minoguchi S, Taniguchi Y, Kato H, Okazaki T, Strobl LJ, Zimber-Strobl U, Bornkamm GW, Honjo T. 1997.** RBP-L, a transcription factor related to RBP-

Jkappa. *Mol. Cell Biol.* **17**:2679-2687.

**Misra UK, Kalita J. 1997.** Movement disorders in Japanese encephalitis. *J Neurol.* **244**:299-303.

**Modis Y, Ogata S, Clements D, Harrison SC. 2004.** Structure of the dengue virus envelope protein after membrane fusion. *Nature.* **427:**313-319.

**Momose F, Basler CF, O'Neill RE, Iwamatsu A, Palese P, Nagata K. 2001.** Cellular splicing factor RAF-2p48/NPI-5/BAT1/UAP56 interacts with the influenza virus nucleoprotein and enhances viral RNA synthesis. *J Virol.* **75**:1899-1908.

**Moslemi AR, Tulinius M, Darin N, Aman P, Holme E, Oldfors A. 2003.** SURF1 gene mutations in three cases with Leigh syndrome and cytochrome c oxidase deficiency. *Neurology.* **61**:991-993.

**Mossman KL, Macgregor PF, Rozmus JJ, Goryachev AB, Edwards AM, Smiley JR. 2001.** Herpes simplex virus triggers and then disarms a host antiviral response. *J Virol.* **75**:750–758.

**Mukhopadhyay S, Kim BS, Chipman PR, Rossmann MG, Kuhn RJ. 2003.** Structure of West Nile virus. *Science*. **302**:248.

**Murphy FA. 1980.** Togavirus morphology and morphogenesis, in *The Togavirus. Biology, Structure, Replication.* Schlesinger RW (ed.) Academic Press, New York. pp 241-316.

**Murphy JC, Fischle W, Verdin E, Sinclair JH. 2002.** Control of cytomegalovirus lytic gene expression by histone acetylation. *EMBO J*. **21**:1112-1120.

**Murray JM, Aaskov JG, Wright PJ. 1993.** Processing of the dengue virus type 2 protein prM and C-prM. *J Gen Virol.* **74**:175-182.

**Nadon R, Shoemaker J. 2002.** Statistical issues with microarrays: processing and analysis. *Trends Genetics.* **18**:265-271.

**Ng ML. 1987.** Ultrastructural studies of Kunjin virus-infected Aedes albopictus cells. *J Gen Virol.* **68**:577-582.

**Ng ML, Howe J, Sreenivasan V, Mulders JJ. 1994.** Flavivirus West Nile (Sarafend) egress at the plasma membrane. *Arch Virol.* **137**:303-313.

**Ng ML, Tan SH, Chu JJH. 2001.** Transport and budding at two distinct sites of visible nucleocapsids of West Nile (Sarafend) virus. *J Med Virol.* **65**:758-764.

**Novick D, Kim SH, Fantuzzi G, Reznikov LL, Dinarello CA, Rubinstein M. 1999.** Interleukin-18 binding protein: a novel modulator of the Th1 cytokine response. *Immunity.* **10**:127-136.

**Nowak T, Farber PM, Wengler G, Wengler G. 1989.** Analysis of the terminal sequences of West Nile virus structural proteins and of the *in vitro* translation of these proteins allow the proposal of a complete scheme of the proteolytic cleavages involved in their synthesis. *Virology.* **169**:365-376.

**Olah J, Orosz F, Keseru GM, Kovari Z, Kovacs J, Hollan S, Ovadi J. 2002.** Triosephosphate isomerase deficiency: a neurodegenerative misfolding disease. *Biochem. Soc. Trans.* **30:**30-38.

**O'Leary DR, Marfin AA, Montgomery SP, Kipp AM, Lehman JA, Biggerstaff BJ, Elko VL, Collins PD, Jones JE, Campbell GL. 2004.** The epidemic of West Nile virus in the United States, 2002. *Vector Borne Zoonotic Dis.* **4**:61-70.

**Omalu BI, Shakir AA, Wang G, Lipkin WI, Wiley CA. 2003.** Fatal fulminant pan-meningo-polioencephalitis due to West Nile virus. *Brain Pathol*. **13:**465–472.

**Paavilainen VO, Bertling E, Falck S, Lappalainen P. 2004.** Regulation of cytoskeletal dynamics by actin-monomer-binding proteins. *Trends Cell Biol.* **14**:386-394.

**Palu E, Liesi P. 2002.** Differential distribution of laminins in Alzheimer disease and

normal human brain tissue. *J Neurosci Res.* **69**:243-256.

**Pan W. 2002.** A comparative review of statistical methods for discovering differentially expressed genes in replicated microarray experiments. *Bioinformatics.* **18**:546-554.

**Papin JF, Vahrson W, Dittmer D. 2004.** SYBR green-based real-time quantitative PCR assay for detection of West Nile virus circumvents false-negative results due to strain variability. *J Clin Microbiol.* **42**:1511-1518.

**Parquet MC, Kumatori A, Hasebe F, Morita K, Igarashi A. 2001.** West Nile virus-induced bax-dependent apoptosis. *FEBS Lett.* **500**:17-24.

**Paulukat J, Bosmann M, Nold M, Garkisch S, Kampfer H, Frank S, Raedle J, Zeuzem S, Pfeilschifter J, Muhl H. 2001.** Expression and release of IL-18 binding protein in response to IFN-gamma. *J Immunol.***167**:7038-7043.

**Petersen LR, Roehrig JT. 2001.** West Nile Virus: a reemerging global pathogen. *Emerg Infect Dis.* **7**:611-14.

**Polentarutti N, Bottazzi B, Di Santo E, Blasi E, Agnello D, Ghezzi P, Introna M, Bartfai T, Richards G, Mantovani A. 2000.** Inducible expression of the long pentraxin PTX3 in the central nervous system. *J Neuroimmunol.* **106**:87-94

**Presti RM, Popkin DL, Connick M, Paetzold S, Virgin HW 4th. 2001.** Novel cell type-specific antiviral mechanism of interferon gamma action in macrophages. *J Exp Med.* **193**:483-496.

**Prosniak M, Hooper DC, Dietzschold B, Koprowski H. 2001.** Effect of rabies virus infection on gene expression in mouse brain. *Proc Natl Acad Sci USA.* **98**:2758–2763.

**Qi B, Qi Y, Watari A, Yoshioka N, Inoue H, Minemoto Y, Yamashita K, Sasagawa T, Yutsudo M. 2003.** Pro-apoptotic ASY/Nogo-B protein associates with ASYIP. *J Cell Physiol.* **196**:312-318.

**Qian N, Frank D, O'Keefe D, Dao D, Zhao L, Yuan L, Wang Q, Keating M, Walsh C, Tycko B. 1997**. The IPL gene on chromosome 11p15.5 is imprinted in humans and mice and is similar to TDAG51, implicated in Fas expression and apoptosis. *Hum Mol Genet.* **6**:2021-2029.

**Qu Q, Sawa H, Suzuki T, Semba S, Henmi C, Okada Y, Tsuda M, Tanaka S, Atwood WJ, Nagashima K. 2004.** Nuclear entry mechanism of the human polyomavirus JC virus-like particle: role of importins and the nuclear pore complex. *J Biol Chem.* **279**:27735-27742.

**Quackenbush J. 2002.** Microarray data normalization and transformation. *Nature Genet Supp.* **32**:496-501.

**Rajeevan MS, Vernon SD, Taysavang N, Unger ER. 2001.** Validation of Array-Based Gene Expression Profiles by Real-Time (Kinetic) RT-PCR. *J Mol Diagn.* **3**:26-31.

**Renne R, Barry C, Dittmer D, Compitello N, Brown PO, Ganem D. 2001.** Modulation of cellular and viral gene expression by the latency-associated nuclear antigen of Kaposi's sarcoma-associated herpesvirus. *J Virol.* **75**:458–468.

**Rey FA, Heinz FX, Mandl C, Kunz C, Harrison SC. 1995** The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. *Nature.* **375**:291–298

**Roehrig JT, Nash D, Maldin B, Labowitz A, Martin DA, Lanciotti RS, Campbell GL. 2003.** Persistence of virus reactive serum immunoglobulin M antibody in confirmed West Nile virus encephalitis cases. *Emerg Infect Dis*. **9**:376–379.

**Rolph MS, Zimmer S, Bottazzi B, Garlanda C, Mantovani A, Hansson GK. 2002.** Production of the long pentraxin PTX3 in advanced atherosclerotic plaques. *Arterioscler Thromb Vasc Biol*. **22**:e10-14.

**Roy A, Yan W, Burns KH, Matzuk MM. 2004**. Tektin3 encodes an evolutionarily conserved putative testicular microtubules-related protein expressed preferentially in male germ cells. *Mol Reprod Dev.* **67**:295-302.

**Rozen S, Skaletsky HJ. 2000.** Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386.

**Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, Quackenbush J. 2003.** TM4: a free, open-source system for microarray data management and analysis. *Biotechniques.* **34**:374-378.

**Saha S, Rangarajan PN. 2003.** Common host genes are activated in mouse brain by Japanese encephalitis and rabies virus. *J Gen Virol.* **84**:1729–1735.

**Salmaggi A, Gelati M, Dufour A, Corsini E, Pagano S, Baccalini R, Ferrero E, Scabini S, Silei V, Ciusani E, De Rossi M. 2002.** Expression and modulation of IFN-gamma-inducible chemokines (IP-10, Mig, and I-TAC) in human brain endothelium and astrocytes: possible relevance for the immune invasion of the central nervous system and the pathogenesis of multiple sclerosis. *J Interferon Cytokine Res.* **22**:631-640.

**Sampson BA, Ambrosi C, Charlot A, Reiber K, Veress JF. 2000.** The pathology of human West Nile virus infection. *Human Pathol.* **31**:527-531.

**Samuel CE. 2001**. Antiviral actions of interferons. *Clin Microbiol Rev.* **14**:778-809.

**Sangster MY, Heliams DB, Mackenzie JS, Shellam GR. 1993.** Genetic studies of flavivirus resistance in inbred strains derived from wild mice: evidence for a new resistance allele at the flavivirus resistance locus (*Flv*). *J Virol.* **67**:340-347.

**Schapira AH. 1998.** Mitochondrial dysfunction in neurodegenerative disorders. *Biochim Biophys Acta.* **1366**:225-233.

**Schena M. 2003.** *Microarray analysis*. NY: Wiley-Liss.

**Seligman SJ, Bucher DJ. 2003**. The importance of being outer: consequences of the distinction between the outer and inner surfaces of flavivirus glycoprotein E. *Trends Microbiol.* **11**:108-110.

**Shaw PJ, Jordan EG. 1995.** The nucleolus. *Annu Rev Cell Dev Biol.* **11**:93-121.

**Shen J, Devery JM, King NJ. 1995.** Early induction of interferon-independent virus-specific ICAM-1 (CD54) expression by flavivirus in quiescent but not proliferating fibroblasts – implications for virus-host interactions. *Virology.* **208**:437-449.

**Shen J, T-To SS, Schrieber L, King NJC. 1997.** Early E-Selectin, VCAM-1, ICAM-1, and late MHC antigen induction on human endothelial cells by flavivirus and comodulation of adhesion molecule expression by immune cytokines. *J Virol.* **71**:9323-9332.

**Shi PY. 2002.** Strategies for the identification of inhibitors of West Nile virus and other flaviviruses. *Curr Opi Invest Drugs.* **3**:1567-1573.

**Shimizu K, Kawabe H, Minami S, Honda T, Takaishi K, Shirataki H, Takai Y. 1996.** SMAP, an Smg GDS-associating protein having arm repeats and phosphorylated by Src tyrosine kinase. *J Biol Chem.* **271**:27013-27017.

**Shin EY, Woo KN, Lee CS, Koo SH, Kim YG, Kim WJ, Bae CD, Chang SI, Kim EG. 2004.** Basic fibroblast growth factor stimulates activation of Rac1 through a p85 betaPIX phosphorylation-dependent pathway. *J Biol Chem.* **279**:1994-2004.

**Shiroguchi K, Ohsugi M, Edamatsu M, Yamamoto T, Toyoshima YY**. **2003.** The second microtubule-binding site of monomeric kid enhances the microtubule affinity. *J Biol Chem*. **278**:22460-22465.

**Shrestha B, Gottlieb D, Diamond MS. 2003.** Infection and injury of neurons by West Nile encephalitis virus. *J Virol.* **77**:13203–13213.

**Shrestha B, Diamond MS. 2004.** Role of CD8+ T cells in control of West Nile virus infection. *J Virol.* **78**:8312-8321.

**Slezak K, Michalik M, Kowalczyk A, Rokita H. 2004**. YY1 is recruited to the cytoplasm of vaccinia virus-infected human macrophages by the Crm1 system. *Virus Res*. **102**:177-184.

**Smithburn KC, Hughes TP, Burke AW, Paul JH. 1940.** A neurotropic virus isolated from the blood of a native of Uganda. *Am J Trop Hyg.* **20**:471-492.

**Solomon TJ, How OM, Beasley DW, Mallewa M. 2003.** West Nile encephalitis. *Brit Med J*. **326**:865–869.

**Solomon TJ, Winter PM. 2004.** Neurovirulence and host factors in flavivirus encephalitis-evidence from clinical epidemiology. *Arch Virol Suppl.* **18**:161-170.

**Soukas A, Cohen P, Socci ND, Friedman JM. 2000.** Leptin-specific patterns of gene expression in white adipose tissue. *Genes Dev.* **14**:963-980.

**Southern EM. 2001.** DNA Microarrays: History and Overview, in *DNA Arrays: Methods and Protocols*, ed. Jang B. Rampal.

**Sriurairatna S, Bhamarapravati N. 1977.** Replication of dengue-2 virus in *Aedes albopictus* mosquitoes. An electron microscopy study. *Am J Trop Med Hyg.* **26**:1199-1205.

**Stadler K, Allison SL, Schalich J, Heinz FX. 1997.** Proteolytic activation of tick-borne encephalitis virus by furin. *J Virol.* **71**:8475-8481.

**Starek M, Babkova H. 1976.** Replication dynamics of tick-borne encephalitis virus in, and glycolytic activity of, human diploid cells. *Acta Virol.* **20**:320-325.

**Steele KE, Linn MJ, Schoepp RJ, Komar N, Geisbert TW, Manduca RM, Calle**

**PP, Rapahel BP, Clippinger TL, Larsen T, Smith J, Lanciotti RS, Panella NA McNamara TS 2000** Pathology of fatal West Nile virus infections in native and exotic birds during the 1999. *Vet Pathol.* **37**:208-224.

**Stone TW. 2002.** Purines and neuroprotection. *Adv Exp Med Biol* **513**:249-280.

**Takeda A, Hamano S, Yamanaka A, Hanada T, Ishibashi T, Mak TW, Yoshimura A, Yoshida H. 2003.** Cutting edge: role of IL-27/WSX-1 signaling for induction of T-bet through activation of STAT1 during initial Th1 commitment. *J Immunol.* **170**:4886-4890.

**Tani K, Mizoguchi T, Iwamatsu A, Hatsuzawa K, Tagaya M. 1999.** p125 is a novel mammalian Sec23p-interacting protein with structural similarity to phospholipid-modifying proteins. *J Biol Chem.* **274**:20505-20512.

**Tseng GC, Oh MK, Rohlin L, Liao JC, Wong WH. 2001.** Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects. *Nucl Acids Res.* **29**:2549-2557.

**Turell MJ, O'Guinn M, Oliver J. 2000.** Potential for New York mosquitoes to transmit West Nile virus. *Am J Trop Med.* **62**:413-414.

**Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. 2002.** Accurate normalization of real time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**:research0034.

**Wang X, Sato R, Brown MS, Hua X, Goldstein JL. 1994.** SREBP-1, a membrane-bound transcription factor released by sterol-regulated proteolysis. *Cell.* **77**:53-62.

**Wang SH, Syu WJ, Huang KJ, Lei HY, Yao CW, King CC, Hu ST. 2002.** Intracellular localization and determination of a nuclear localization signal of the core protein of dengue virus. *J Gen Virol.* **83**:3093-3102.

**Wang Y, Lobigs M, Lee E, Meullbacher A. 2003.** CD8+ T cells mediate recovery

and immunopathology in West Nile virus encephalitis. *J Virol.* **77**:13323–13334.

**Wang X, Seed B. 2003.** A PCR primer bank for quantitative gene expression analysis. *Nucleic Acids Res.* **31**:1-8.

**Warke RV, Xhaja K, Martin KJ, Fournier MF, Shaw SK, Brizuela N, De Bosch N, Lapointe D, Ennis FA., Rothman AL, Bosch I. 2003.** Dengue virus induces novel changes in gene expression of human umbilical vein endothelial cells. *J Virol.* **77**:11822-11832.

**Warrington JA, Dee S, Trulson M. 2000**. Large-scale genomic analysis using Affymetrix GeneChip probe arrays, in *Microarray Biochip Technology*, ed. Mark Schena. Natick, MA: Eaton.

**Wen L, Wong J, Penfold PL. 2001.** Changes in transcellular resistance in retinal pigment epithelium and endothelium after flavivirus infection correlate with changes in distribution of ZO-1 and β-catenin [Abstract]. *Sixth International Symposium on Positive Strand RNA Viruses.* Paris, France.

**Wengler G. 1989.** Cell-associated West Nile flavivirus is covered with E+ pre-M protein heterodimers which are destroyed and reorganized by proteolytic cleavage during virus release. *J Virol.* **63**:2521-2526.

**Wengler G, Wengler G, Rey FA. 1999.** The isolation of the ectodomain of the alphavirus E1 protein as a soluble hemagglutinin and its crystallization. *Virology.* **257**:472-482.

**Westaway EG, Ng ML. 1980.** Replication of flaviviruses: Separation of membrane translation sites of Kunjin virus proteins and of cell proteins. *Virology.* **106**:107-122.

**Westaway EG, Mackenzie JM, Khromykh AA. 2002.** Replication and gene function in Kunjin virus. *Curr Top Microbiol Immunol*. **267**:323-351.

**Wonsey DR, Zeller KI, Dang CV. 2002.** The c-Myc target gene PRDX3 is required

for mitochondrial homeostasis and neoplastic transformation. *Proc Natl Acad Sci USA.* **99**:6649-6654.

**Worley J. 2000.** *Microarray Biochip Technology* (Schena M, ed.), Eaton Publishing/Biotechniques Books, Natick, MA, pp. 65-85.

**Wyman AH, Chi M, Riley J, Carayannopoulos MO, Yang C, Coker KJ, Pessin JE, Moley KH. 2003.** Syntaxin 4 expression affects glucose transporter 8 translocation and embryo survival. *Mol Endocrinol.* **17**:2096-2102.

**Xiao H, Neuveut C, Benkirane M, Jean KT. 1998.** Interaction of the second coding exon of Tat with human EF-1 delta delineates a mechanism for HIV-1-mediated shut-off of host mRNA translation. *Biochem Biophys Res Commun.* **244**:384-389.

**Xiao SY, Guzman H, Zhang H, Travassos da Rosa AP, Tesh RB. 2001.** West Nile virus infection in the golden hamster (Messocricetus auratus): a model for West Nile encephalitis. *Emerg Infect Dis*. **7**:714–721.

**Yang JS, Ramanathan MP, Muthumani K, Choo AY, Jin SH, Yu QC, Hwang DS, Choo DK, Lee MD, Dang K, Tang W, Kim JJ, Weiner DB. 2002a.** Induction of inflammation by West Nile virus capsid through the caspase-9 apoptotic pathway. *Emerg Infect Dis.* **8**:1379-1384

**Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, Speed TP. 2002b.** Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res*. **30**:e15.

**Yang IV, Chen E, Hasseman JP, Liang W, Frank BC, Wang S, Sharov V, Saeed AI, White J, Li J, Lee NH, Yeatman TJ, Quackenbush J. 2002c.**Within the fold: assessing differential expression measures and reproducibility in microarray assays. *Genome Biol.* **3**, research0062.1–0062.12.

**Yeung KY, Haynor DR, Ruzzo WL. 2001.** Validating clustering for gene expression

data. *Bioinformatics.* **17**:309-318.

**Zagursky RJ, Russell D. 2001.** Bioinformatics: use in bacterial vaccine discovery. *Biotechniques* **31**:636–659.

**Zhang W, Chipman PR, Corver J, Johnson PR, Zhang Y, Mukhopadhyay S, Baker TS, Strauss JH, Rossmann MG, Kuhn RJ. 2003.** Visualization of membrane protein domains by cryo-electron microscopy of dengue virus. *Nat Struct Biol.* **10**:907-912.

**Zhu H, Cong JP, Mamtora G, Gingeras T, Shenk T. 1998.** Cellular gene expression altered by human cytomegalovirus: global monitoring with oligonucleotide arrays. *Proc Natl Acad Sci USA.* **95**:14470–14475.

## Appendix 1: Media for Tissue Culture of Cell Lines

**a) Dulbecco's Modified Eagle's Medium (DMEM): Growth Medium for HeLa and A172 Cells**

| Item | Amount | Source |
|---|---|---|
| DMEM Powder | 17.3 g | Sigma, USA |
| $NaHCO_3$ | 2.2 g | Merck, Germany |
| Foetal Calf Serum (FCS) | 100 ml | Hyclone, Switzerland |
| Type 1 Reagent Grade Water (NANOpure) | 900 ml | Barnstead, USA |

1 bottle of DMEM powder was dissolved in 900 ml of NANOpure water. 2.2 g of $NaHCO_3$ was then added and the pH adjusted to 7.2. Sterilization of media was carried out by filtration through a 0.22 μm filter (Sterivex G-S, Millipore, USA). The media was then stored in aliquots at 4°C. FCS was added as required before use.

**b) Medium 199: Growth Medium for Vero Cells**

| Item | Amount | Source |
|---|---|---|
| M199 Powder | 11.0 g | Sigma, USA |
| $NaHCO_3$ | 2.2 g | Merck, Germany |
| Foetal Calf Serum (FCS) | 100 ml | Hyclone, Switzerland |
| Type 1 Reagent Grade Water (NANOpure) | 900 ml | Barnstead, USA |

1 bottle of DMEM powder was dissolved in 900 ml of NANOpure water. 2.2 g of $NaHCO_3$ was then added and the pH adjusted to 7.2. Sterilization of media was carried out by filtration through a 0.22 μm filter (Sterivex G-S, Millipore, USA). The media was then stored in aliquots at 4°C. FCS was added as required before use.

**c) 10X Phosphate Buffered Saline [(PBS) (pH7.4)]**

| Item | Amount | Source |
|---|---|---|
| NaCl | 80.0 g | Merck, Germany |
| KCl | 2.0 g | Merck, Germany |
| $KH_2PO_4$ | 2.0 g | Merck, Germany |
| $Na_2HPO_4$ | 11.5 g | Merck, Germany |
| Type 1 Reagent Grade Water (NANOpure) | 1000 ml | Barnstead, USA |

To prepare 10X PBS stock solution, the specified items were added to 900 ml of NANOpure water and the pH adjusted to 7.4. The solution was topped up to 1 litre and then autoclaved at 121°C for 15 min. The stock solution was then stored at room temperature. To prepare the working 1X PBS, the stock solution was diluted 1:10 with NANOpure water. The solution was then autoclaved at 121°C for 15 min and subsequently stored at 4°C.

**d) Trypsin/Versene (ATV) Solution**

| Item | Amount | Source |
|---|---|---|
| NaCl | 80.0 g | Merck, Germany |
| KCl | 4.0 g | Merck, Germany |
| D-glucose | 10.0 g | Analar, UK |
| $NaHCO_3$ | 5.8 g | Merck, Germany |
| Trypsin | 5.0 g | Difco, New Zealand |
| Versene (EDTA) | 2 g | Sigma, USA |
| Type 1 Reagent Grade Water (NANOpure) | 100 ml | Barnstead, USA |

The items were added to 100 ml of NANOpure water and the mixture was heated to 30°C with occasional shaking (for 3-4 hr) to dissolve the trypsin. The solution was then filtered through a 22 μm filter (Sterivex G-S, Millipore, USA). Ten ml of this solution was added to 90 ml of sterile NANOpure water to give 1X ATV solution.

## Appendix 2: Reagents for Plaque Assay

**a) Virus Diluent**

| Item | Amount | Source |
|---|---|---|
| Hank's Balanced Salt Solution | 11.0 g | Sigma, USA |
| 10% Bovine Serum Albumin | 1 ml | CSL, Australia |
| $NaHCO_3$ | 2.2 g | Merck, Germany |
| Type 1 Reagent Grade Water (NANOpure) | 99 ml | Barnstead, USA |

All items were dissolved in the water to give 100 ml. The solution was sterilized by filtration through a 0.22 μm filter unit (Sterivex G-S, Millipore, USA). Storage was at 4°C.

**b) Maintenance Medium (DMEM) for HeLa and A172 Cells**

| Item | Amount | Source |
|---|---|---|
| DMEM Powder | 17.3 g | Sigma, USA |
| Foetal Calf Serum | 20 ml | Hyclone, Switzerland |
| $NaHCO_3$ | 2.2 g | Merck, Germany |
| Type 1 Reagent Grade Water (NANOpure) | 980 ml | Barnstead, USA |

1 bottle of DMEM powder was dissolved in 980 ml of water, and 2.2 g of $NaHCO_3$ was then added. The pH was adjusted to 7.2 and sterilization was carried out by filtration through a 0.22 μm filter (Sterivex G-S, Millipore, USA). The media was then stored in aliquots at 4°C. FCS was added as required prior to use.

**c) Maintenance Medium (M199) for Vero Cells**

| **Item** | **Amount** | **Source** |
|---|---|---|
| M199 Powder | 11.0 g | Sigma, USA |
| Foetal Calf Serum | 20 ml | Hyclone, Switzerland |
| $NaHCO_3$ | 2.2 g | Merck, Germany |
| Type 1 Reagent Grade Water (NANOpure) | 980 ml | Barnstead, USA |

1 bottle of M199 powder was dissolved in 980 ml of water, and 2.2 g of $NaHCO_3$ was then added. The pH was adjusted to 7.2 and sterilization was carried out by filtration through a 0.22 μm filter (Sterivex G-S, Millipore, USA). The media was then stored in aliquots at 4°C. FCS was added as required prior to use.

**d) Overlay Medium**

| **Item** | **Amount** | **Source** |
|---|---|---|
| M199 Powder | 11.0 g | Sigma, USA |
| Foetal Calf Serum | 40 ml | Hyclone, Switzerland |
| $NaHCO_3$ | 4.4 g | Merck, Germany |
| Type 1 Reagent Grade Water (NANOpure) | 460 ml | Barnstead, USA |

The reagents were dissolved in the water to produce a 2X concentrate nutrient medium. Two hundred ml of this 2X concentrate medium was added to an equal volume of 2.5% carboxymethyl-cellulose solution (See below. Appendix 2f).

**e) 1% Crystal Violet Staining Solution**

| Item | Amount | Source |
|---|---|---|
| Crystal Violet Powder | 1.85 g | BDH, UK |
| 37% Formaldehyde | 250 ml | Merck, Germany |
| PBS | 120 ml | see Appendix 1(c) |

**f) 2.5% Carboxymethyl-Cellulose Solution**

| Item | Amount | Source |
|---|---|---|
| Sodium salt of Carboxymethyl-Cellulose | 2.5 g | CalBiochem, USA |
| Type 1 Reagent Grade Water (NANOpure) | 100 ml | Barnstead, USA |

2.5 g of sodium salt of carboxymethyl-cellulose (Aquacide II, CalBiochem, USA) was added to 100 ml of the water. The solution was mixed, autoclaved for 15 min at 121°C and stored at 4°C.

## Appendix 3: Reagents for Genomic Expression Studies

### a) Diethylpyrocarbonate (DEPC) Treated dd$H_2O$

| Item | Amount | Source |
|---|---|---|
| Diethylpyrocarbonate (DEPC) | 1.0 ml | Sigma, USA |
| Type 1 Reagent Grade Water (NANOpure) | 999 ml | Barnstead, USA |

To make 1 litre of DEPC treated (RNase-free) water, 1ml of DEPC was added to 999 ml of water. The treated water was incubated overnight at room temperature with stirring. The water was then autoclaved at 121°C for 15 min.

### b) Formaldehyde-Agarose (FA) Gel

| Item | Amount | Source |
|---|---|---|
| Agarose | 1.2 g | Merck, Germany |
| 10x FA Gel Buffer | 10 ml | see Appendix 3i |
| 37% (12.3M) Formaldehyde | 1.8 ml | Sigma, USA |
| Ethidium Bromide (10µg/ml) | 1 µl | see Appendix 3j |
| DEPC Water | 90 ml | see Appendix 3a |

The agarose and 10x FA gel buffer was mixed and topped up to 100 ml with DEPC water. It was then swirled and microwaved until completely dissolved. Formaldehyde and ethidium bromide was then added and mixed at 60°C.

**c) 1x Formaldehyde-Agarose (FA) Gel Running Buffer**

| Item | Amount | Source |
|---|---|---|
| 10x FA Gel Buffer | 100 ml | see Appendix 3i |
| 37% (12.3M) Formaldehyde | 20 ml | Sigma, USA |
| DEPC Water | 880 ml | see Appendix 3a |

Store at room temperature after preparation.

**d) 5x RNA Loading Buffer**

| Item | Amount | Source |
|---|---|---|
| Bromophenol Blue | 16 µl | Sigma, USA |
| Formamide | 3.084 ml | Sigma, USA |
| 37% (12.3M) Formaldehyde | 720 µl | Sigma, USA |
| 100% Glycerol | 2 ml | Merck, Germany |
| 500mM EDTA, pH8.0 | 80 µl | Merck, Germany |
| 10x FA Gel Buffer | 4 ml | See Appendix 3i |
| DEPC Water | to 10 ml | see Appendix 3a |

Store at -20°C after preparation.

**e) 2.5M NaOH**

| Item | Amount | Source |
|---|---|---|
| NaOH | 20.0 g | Merck, Germany |
| DEPC Water | 200 ml | see Appendix 3a |

**f) 2M HEPES free acid**

| Item | Amount | Source |
|---|---|---|
| 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES) | 95.3 g | Merck, Germany |
| DEPC Water | 200 ml | see Appendix 3a |

**g) Wash Solution 1**

| Item | Amount | Source |
|---|---|---|
| 20x SSC | 300 ml | Sigma, USA |
| Triton X-102 | 0.5 ml | Supplied with kit. |
| DEPC Water | 700 ml | see Appendix 3a |

Mix the items in a nuclease-free graduated cylinder. Pass it through a 0.22 μm filter (Sterivex G-S, Millipore, USA) and store at room temperature.

**h) Wash Solution 2**

| Item | Amount | Source |
|---|---|---|
| 20x SSC | 5 ml | Sigma, USA |
| 10% Triton X-102 | 0.5 ml | Supplied with kit. |
| DEPC Water | 995 ml | see Appendix 3a |

Mix the items in a nuclease-free graduated cylinder. Pass it through a 0.22 μm filter (Sterivex G-S, Millipore, USA) and store at room temperature.

**i) 10x FA Gel Buffer**

| Item | Amount | Source |
|---|---|---|
| 200mM 3-[N-morpholino]propanesulfonic acid (MOPS) (free acid) | 41.86 g | Sigma, USA |
| 50mM Sodium acetate | 4.1 g | Merck, Germany |
| 10mM EDTA | 2.9 g | Merck, Germany |
| DEPC Water | to 1000 ml | see Appendix 3a |

Mix the above items with a magnetic stirrer. pH to 7.0 with NaOH.

**j) Ethidium Bromide (EtBr)**

| Item | Amount | Source |
|---|---|---|
| Ethidium Bromide (EtBr) | 0.2 g | Sigma, USA |
| DEPC Water | 20 ml | see Appendix 3a |

EtBr was dissolved in 20 ml of DEPC water by stirring with a magnetic stirrer at room temperature for several hours. The solution was then stored in the dark at 4°C.

## Appendix 4: Reagents for Immunofluorescence

**a) 10% Bovine Serum Albumin (BSA)**

| Item | Amount | Source |
|---|---|---|
| Bovine serum albumin powder, Fraction V | 10 g | Gibco BRL, USA |
| PBS | 100 ml | see Appendix 1c |
| Glycerol | 9 ml | Merck, Germany |

The solution was aliquoted and stored at –20°C. To obtain 0.1% BSA, 1 ml of 10% BSA was diluted in 100 ml of PBS.

**b) Mountant**

| Item | Amount | Source |
|---|---|---|
| 1,4-Diazabicyclo[2.2.2]octane (DABCO) | 1 g | Sigma, USA |
| 0.1 M Tris-HCl (pH 7.6) | 1 ml | Merck, Germany |
| Glycerol | 9 ml | Merck, Germany |

The mountant was aliquoted and then stored at –20°C.

## Appendix 5: List of Oligonucleotides

| Gene | Symbol | Sequence (5'-3') |
|---|---|---|
| West Nile (Sarafend) virus envelope protein | WN(S)E (+) | TCCCTGAACGACCTTACACC |
| | WN(S)E (-) | ACTAAACGGGGGTTCGAGTT |
| Hypoxanthine phosphoribosyl-transferase 1 | HPRT1 (+) | TGACACTGGCAAAACAATGCA |
| | HPRT1 (-) | GGTCCTTTTCACCAGCAAGCT |
| IFN induced transmembrane protein 1 | IFITM1 (+) | TTCATCCTGTCACTGGTATTCG |
| | IFITM1 (-) | AGGTGTGTGGGTATAAACTGCT |
| Tissue factor pathway inhibitor 2 | TFPI2 (+) | TTGAGAACAGGTTTCCAGATGA |
| | TFPI2 (-) | AGGTGAAAGCATCACAGGTTCT |
| DnaJ (Hsp40) homolog, subfamily A, member 1 | DNAJA1 (+) | GAAGGCATGGACATTGATGACC |
| | DNAJA1 (-) | TGGGGGATCTTGCTTCTTTCG |
| Dual specificity phosphatase 1 | DUSP1 (+) | CTTAGCGTCAAGACATTTGCTG |
| | DUSP1 (-) | TGTGTCGTCGGGAATAATACTG |
| Ras homolog gene family, member I | ARHI (+) | CCTGTAGCGTGTAGAAAACGTG |
| | ARHI (-) | CTCCTCTTCAAATGCCAATGTT |
| 14-3-3 zeta | YWHAZ (+) | TGGTACATTGTGGCTTCAAAAG |
| | YWHAZ (-) | CTACACCTGTGACTGGAACCAA |
| Hypothetical protein FLJ13855 | FLJ13855 (+) | CCTGGCTTTGTTGCTTCTCTAT |
| | FLJ13855 (-) | TAAATCTGGAAGTACGGGGAGA |
| Cyclin-E binding protein 1 | HERC5 (+) | TCAACAACAACAGAGGATTTGG |
| | HERC5 (-) | CTTCATCCCTCTTTCCTGCTTA |
| 40S ribosomal protein S12 | RPS12 (+) | TGTTCTGACTGCTGACAATTCC |
| | RPS12 (-) | GGCAGTCTTCAGAACCTCTTGT |
| Ribosomal protein L7 | RPL7 (+) | TAAGAAGCGAATTGCTTTGACA |
| | RPL7 (-) | AAATGGGTGGTCTTTTTCTTCA |

# Appendix 6: List of Differentially Regulated Genes in A172 Cells at 24 h Post-Infection

| GB Acc | Fold Change | Gene Name |
|---|---|---|
| J04164 | 12.00 | interferon induced transmembrane protein 1 (9-27) |
| X03557 | 10.74 | interferon-induced protein with tetratricopeptide repeats 1 |
| M13755 | 9.50 | interferon, alpha-inducible protein (clone IFI-15K) |
| AF095154 | 6.42 | C1q-related factor |
| BC007922 | 5.57 | interferon stimulated gene 20kDa |
| BC005330 | 5.34 | tissue factor pathway inhibitor 2 |
| X52541 | 4.79 | early growth response 1 |
| AF038963 | 4.19 | DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide |
| BC002601 | 4.13 | nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha |
| X67325 | 4.03 | interferon, alpha-inducible protein 27 |
| M14660 | 3.76 | interferon-induced protein with tetratricopeptide repeats 2 |
| AF080246 | 3.53 | chromosome 22 open reading frame 1 |
| AJ225089 | 3.46 | 2'-5'-oligoadenylate synthetase-like |
| X63613 | 3.44 | pentaxin-related gene, rapidly induced by IL-1 beta |
| BC027882 | 3.38 | indoleamine-pyrrole 2,3 dioxygenase |
| M30818 | 3.17 | myxovirus (influenza virus) resistance 2 (mouse) |
| X57351 | 3.04 | interferon induced transmembrane protein 2 (1-8D) |
| U23070 | 2.97 | putative transmembrane protein |
| AB006746 | 2.94 | phospholipid scramblase 1 |
| V00546 | 2.79 | interferon, beta 1, fibroblast |
| AK002148 | 2.75 | hypothetical protein FLJ11286 |
| AF061738 | 2.73 | leucine aminopeptidase 3 |
| AF134053 | 2.62 | Kruppel-like factor 2 (lung) |
| BC009353 | 2.50 | HL-60-induced differentiation immediate-early protein |
| AK026672 | 2.50 | hypothetical protein FLJ39885 |
| AF288815 | 2.47 | guanylate binding protein 5 |
| AB024964 | 2.43 | recombining binding protein suppressor of hairless (Drosophila)-like |
| BC013292 | 2.33 | 5'-nucleotidase, cytosolic III |
| AB027289 | 2.33 | cyclin-E binding protein 1 |
| AF063613 | 2.32 | 2'-5'-oligoadenylate synthetase 3, 100kDa |
| BC012532 | 2.31 | chemokine (C-X-C motif) ligand 11 |
| X55019 | 2.30 | cholinergic receptor, nicotinic, delta polypeptide |
| M62800 | 2.29 | Sjogren syndrome antigen A1 (52kDa, ribonucleoprotein autoantigen SS-A/Ro) |
| AF070674 | 2.24 | baculoviral IAP repeat-containing 3 |
| AF311913 | 2.21 | potassium voltage-gated channel, subfamily H (eag-related), member 6 |
| BC004489 | 2.20 | major histocompatibility complex, class I, C |
| AB023206 | 2.20 | angiomotin like 2 |
| AF064090 | 2.19 | tumor necrosis factor (ligand) superfamily, member 14 |
| M77693 | 2.18 | spermidine/spermine N1-acetyltransferase |
| BC005010 | 2.16 | zinc finger protein 36, C3H type-like 2 |
| X07834 | 2.12 | superoxide dismutase 2, mitochondrial |
| AF376770 | 2.12 | Member of the gamma-glutamyltranspeptidase family |
| U41315 | 2.11 | makorin, ring finger protein, 4 |
| AF442151 | 2.11 | viperin |
| BC010954 | 2.10 | chemokine (C-X-C motif) ligand 10 |
| BC015692 | 2.09 | Protein containing two PDZ, DHR, or GLGF domains |
| X67330 | 2.09 | Protein of unknown function |

| | | |
|---|---|---|
| BC016648 | 2.08 | FOS-like antigen 1 |
| AL713764 | 2.06 | tissue inhibitor of metalloproteinase 2 |
| BC013161 | 2.06 | 28kD interferon responsive protein |
| AK027372 | 2.05 | hypothetical protein FLJ14466 |
| BC015890 | 2.05 | hypothetical protein FLJ13855 |
| AK001176 | 2.04 | Protein of unknown function |
| AK000707 | 2.03 | hypothetical protein FLJ20700 |
| AB017335 | 2.02 | MYC-associated zinc finger protein (purine-binding transcription factor) |
| U19261 | 2.01 | TNF receptor-associated factor 1 |
| AJ315767 | 2.00 | O-linked N-acetylglucosamine (GlcNAc) transferase |
| X83218 | -2.00 | ATP synthase, H+ transporting, mitochondrial F1 complex, O subunit |
| AF283772 | -2.01 | ribosomal protein L15 |
| BC007850 | -2.02 | solute carrier family 25 |
| J03459 | -2.02 | leukotriene A4 hydrolase |
| BC017820 | -2.03 | ribosomal protein L41 |
| M36072 | -2.03 | ribosomal protein L7a |
| BC030135 | -2.04 | SH3 domain binding glutamic acid-rich protein like 3 |
| AJ132583 | -2.04 | aminopeptidase puromycin sensitive |
| U15008 | -2.07 | small nuclear ribonucleoprotein D2 polypeptide 16.5kDa |
| AF348700 | -2.09 | ubiquitin A-52 residue ribosomal protein fusion product 1 |
| BC010004 | -2.10 | farnesyl diphosphate synthase |
| U73824 | -2.11 | eukaryotic translation initiation factor 4 gamma, 2 |
| BC005230 | -2.11 | ubiquinol-cytochrome c reductase binding protein |
| AF217963 | -2.11 | melanoma antigen, family D, 1 |
| BC001178 | -2.11 | ATP synthase, H+ transporting, mitochondrial F0 complex, subunit F6 |
| AK000373 | -2.11 | hypothetical protein FLJ20366 |
| AJ251760 | -2.11 | GNAS complex locus |
| AK000575 | -2.12 | host cell factor C1 regulator 1 (XPO1 dependant) |
| BC017360 | -2.12 | ribosomal protein L26-like 1 |
| BC005338 | -2.12 | capping protein (actin filament) muscle Z-line, alpha 2 |
| AB011173 | -2.12 | KIAA0601 protein |
| S65738 | -2.13 | destrin (actin depolymerizing factor) |
| BC006229 | -2.13 | cytochrome c oxidase subunit Vb |
| BC012460 | -2.14 | LIM and SH3 protein 1 |
| AB011159 | -2.14 | NCK-associated protein 1 |
| BC003655 | -2.16 | ribosomal protein, large, P0 |
| BC003409 | -2.17 | ovarian carcinoma immunoreactive antigen |
| M10036 | -2.17 | triosephosphate isomerase 1 |
| BC016512 | -2.17 | ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide |
| AY034001 | -2.17 | nascent-polypeptide-associated complex alpha polypeptide |
| S73591 | -2.17 | thioredoxin interacting protein |
| BC004975 | -2.17 | cyclin I |
| X07696 | -2.18 | keratin 15 |
| BC022510 | -2.20 | ADP-ribosylation factor domain protein 1, 64kDa |
| AF020038 | -2.20 | isocitrate dehydrogenase 1 (NADP+), soluble |
| AF035737 | -2.20 | general transcription factor II, i |
| D14710 | -2.21 | ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit, isoform 1 |
| BC015887 | -2.21 | WD repeat and SOCS box containing protein 2 |
| Z14244 | -2.21 | cytochrome c oxidase subunit VIIb |
| BC017100 | -2.22 | ring finger protein 130 |
| M83248 | -2.22 | secreted phosphoprotein 1 |
| BC030828 | -2.22 | EGF-like repeats and discoidin I-like domains 3 |

| | | |
|---|---|---|
| U00968 | -2.23 | sterol regulatory element binding transcription factor 1 |
| BC010114 | -2.23 | ribosomal protein L23 |
| AF283772 | -2.25 | ribosomal protein L15 |
| AL136693 | -2.25 | cytochrome b reductase 1 |
| Z49861 | -2.25 | histone 3, H3 |
| BC007062 | -2.27 | peroxiredoxin 3 |
| AJ242586 | -2.28 | arginyl aminopeptidase (aminopeptidase B) |
| D50369 | -2.28 | Ubiquinone binding protein |
| AK027426 | -2.29 | signal peptidase 12kDa |
| BC017396 | -2.30 | prostatic binding protein |
| BC007748 | -2.30 | ribosomal protein L4 |
| BC016325 | -2.30 | ADP-ribosylation factor 4 |
| D49737 | -2.31 | succinate dehydrogenase complex, subunit C, integral membrane protein, 15kDa |
| BC015041 | -2.36 | vesicle amine transport protein 1 homolog (T californica) |
| AF169650 | -2.36 | High mobility group protein 1 |
| X14420 | -2.36 | collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant) |
| U88895 | -2.37 | Putative protein encoded by human endogenous retrovirus H (HERV-H) |
| AF109873 | -2.39 | CREBBP/EP300 inhibitory protein 1 |
| BC001015 | -2.41 | cytochrome c oxidase subunit VIb |
| BC002964 | -2.42 | profilin 2 |
| BC005143 | -2.42 | cytoskeleton related vitamin A responsive protein |
| BC006262 | -2.42 | Protein containing a reeler domain |
| AF037364 | -2.43 | paraneoplastic antigen MA1 |
| BC016350 | -2.43 | ATP synthase, H+ transporting, mitochondrial F0 complex, subunit b, isoform 1 |
| AL132665 | -2.49 | BCL2/adenovirus E1B 19kDa interacting protein 3-like |
| BC010514 | -2.50 | clusterin |
| AB053293 | -2.50 | Rho GTPase activating protein 18 |
| BC025372 | -2.50 | calponin 3, acidic |
| M31520 | -2.51 | ribosomal protein S24 |
| BC000614 | -2.52 | histone deacetylase 3 |
| BC017875 | -2.52 | Cyclophilin A |
| BC016300 | -2.53 | S100 calcium binding protein A4 |
| M82809 | -2.53 | annexin A4 |
| Z11793 | -2.54 | selenoprotein P, plasma, 1 |
| AB062450 | -2.55 | NIMA (never in mitosis gene a)-related kinase 7 |
| M61199 | -2.55 | sperm specific antigen 2 |
| Z11692 | -2.60 | eukaryotic translation elongation factor 2 |
| X76534 | -2.63 | glycoprotein (transmembrane) nmb |
| AJ277724 | -2.64 | histone deacetylase 8 |
| AL080089 | -2.64 | ATP synthase, H+ transporting, mitochondrial F0 complex, subunit c, isoform 1 |
| AK000265 | -2.68 | EPS8-like 1 |
| U41766 | -2.69 | a disintegrin and metalloproteinase domain 9 (meltrin gamma) |
| U96750 | -2.73 | ras homolog gene family, member I |
| AF124993 | -2.74 | peroxiredoxin 5 |
| AJ330246 | -2.77 | Member of the galactosyltransferase family |
| BC003352 | -2.79 | tumor protein, translationally-controlled 1 |
| X78669 | -2.84 | reticulocalbin 2, EF-hand calcium binding domain |
| AJ132694 | -2.87 | ras-related C3 botulinum toxin substrate 1 |
| BC000490 | -2.88 | eukaryotic translation initiation factor 3, subunit 5 epsilon, 47kDa |
| M90354 | -2.93 | basic transcription factor 3, like 1 |
| D84295 | -2.96 | tetratricopeptide repeat domain 3 |
| U14966 | -2.97 | ribosomal protein L5 |

| | | |
|---|---|---|
| AF100615 | -2.97 | mortality factor 4 like 1 |
| Y13286 | -2.97 | GDP dissociation inhibitor 2 |
| X57347 | -2.99 | tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein |
| D00943 | -3.00 | myosin, heavy polypeptide 6, cardiac muscle, alpha |
| M84711 | -3.04 | ribosomal protein S3A |
| Z83737 | -3.06 | histone 1, H3j |
| BC008850 | -3.10 | ribosomal protein L7 |
| M60746 | -3.18 | histone 1, H3e |
| BC008443 | -3.27 | signal recognition particle 9kDa |
| BC031012 | -3.29 | eukaryotic translation elongation factor 1 gamma |
| BC001708 | -3.33 | ribosomal protein S3A |
| BC003623 | -3.33 | tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein |
| BC017701 | -3.40 | hypothetical protein DKFZp434L142 |
| AF348078 | -3.48 | G protein-coupled receptor 91 |
| AK025643 | -3.50 | Ribosomal protein S12 |
| U96876 | -3.70 | insulin induced gene 1 |
| BC000931 | -3.82 | ATP synthase, H+ transporting, mitochondrial F1 complex, gamma polypeptide 1 |
| X83550 | -3.95 | histone 1, H3i |
| BC007552 | -4.15 | H3 histone family 3B |
| AK025927 | -4.33 | hypothetical protein MGC8721 |
| AF023860 | -4.47 | Protein with strong similarity to cyclophilin A |
| BC030690 | -4.50 | sema domain, immunoglobulin domain (Ig) |

## Appendix 7: List of genes that were constantly upregulated during WNV infection. (GpI)

| GB Acc | Symbol | Gene Name |
|---|---|---|
| ***Immune response*** | | |
| X03557 | IFIT1 | interferon-induced protein with tetratricopeptide repeats 1 |
| M14660 | IFIT2 | interferon-induced protein with tetratricopeptide repeats 2 |
| J04164 | IFITM1 | interferon induced transmembrane protein 1 (9-27) |
| X67325 | IFI27 | interferon, alpha-inducible protein 27 |
| AJ225089 | OASL | 2'-5'-oligoadenylate synthetase-like |
| AF110801 | IL18BP | interleukin 18 binding protein |
| BC011171 | SERPING1 | serine (or cysteine) proteinase inhibitor, clade G (C1 inhibitor), member 1, (angioedema, hereditary) |
| BC032589 | B2M | beta-2-microglobulin |
| M28827 | CD1C | CD1C antigen, c polypeptide |
| Y00636 | CD58 | CD58 antigen, (lymphocyte function-associated antigen 3) |
| M30818 | MX2 | myxovirus (influenza virus) resistance 2 (mouse) |
| ***Immune response: inflammation*** | | |
| X63613 | PTX3 | pentaxin-related gene, rapidly induced by IL-1 beta |
| U32849 | NMI | N-myc (and STAT) interactor |
| AJ238044 | BDKRB1 | bradykinin receptor B1 |
| BC014451 | PROCR | protein C receptor, endothelial (EPCR) |
| ***Others: organogenesis*** | | |
| AF022654 | SHOX2 | short stature homeobox 2 |
| BC009983 | MAB21L2 | mab-21-like 2 (C. elegans) |
| U05237 | FALZ | fetal Alzheimer antigen |
| U41901 | LSAMP | limbic system-associated membrane protein |
| X73902 | LAMC2 | laminin, gamma 2 |
| ***Others: signal transduction*** | | |
| AK027780 | GPR34 | G protein-coupled receptor 34 |
| BC002601 | NFKBIA | nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha |
| BC012513 | ARHE | ras homolog gene family, member E |
| U37547 | BIRC2 | baculoviral IAP repeat-containing 2 (apoptosis inhibitor) |
| ***Others*** | | |
| AB014515 | N4BP1 | Nedd4 binding protein 1 |
| AB027289 | LOC51191 | cyclin-E binding protein 1 |
| AF099011 | EHD1 | EH-domain containing 1 |
| AJ131890 | POLL | polymerase (DNA directed), lambda |
| AK000146 | FLJ20139 | CGI-30 protein |
| AK000644 | FLJ20637 | hypothetical protein FLJ20637 |
| AK021416 | LGP2 | likely ortholog of mouse D11lgp2 (DEAD box) |
| AK024401 | ABCB8 | ATP-binding cassette, sub-family B (MDR/TAP), member 8 |
| AK025485 | C10orf2 | progressive external ophthalmoplegia 1 |
| AK026672 | FLJ39885 | hypothetical protein FLJ39885 |
| AL080121 | DKFZP564O0823 | DKFZP564O0823 protein |
| AL512725 | MIDN | midnolin |
| BC001946 | BRUNOL4 | bruno-like 4, RNA binding protein |
| BC005361 | PSMA4 | proteasome (prosome, macropain) subunit, alpha type, 4 |
| BC006392 | CNP | 2',3'-cyclic nucleotide 3' phosphodiesterase |
| BC013161 | IFRG28 | 28kD interferon responsive protein |
| BC029162 | MGC35440 | hypothetical protein MGC35440 |
| J04056 | CBR1 | carbonyl reductase 1 |
| L13852 | UBE1L | ubiquitin-activating enzyme E1-like |
| M77693 | SAT | spermidine/spermine N1-acetyltransferase |

## Appendix 8: List of genes that were constantly downregulated during WNV infection. (GpII)

| GB Acc | Symbol | Gene Name |
|---|---|---|
| ***Cell proliferation*** | | |
| D16431 | HDGF | hepatoma-derived growth factor (high-mobility group protein 1-like) |
| S67334 | PIK3CB | phosphoinositide-3-kinase, catalytic, beta polypeptide |
| X13916 | LRP1 | low density lipoprotein-related protein 1 (alpha-2-macroglobulin receptor) |
| ***Transport*** | | |
| AF151878 | LOC51644 | coatomer protein complex, subunit zeta 1 ( clathrin adaptor complex) |
| ***Transcription*** | | |
| AB007874 | ZNF297B | zinc finger protein 297B |
| AK027559 | SIN3A | SIN3 homolog A, transcriptional regulator (yeast) |
| ***Signal transduction*** | | |
| BC008770 | GPR56 | G protein-coupled receptor 56 |
| S67334 | PIK3CB | phosphoinositide-3-kinase, catalytic, beta polypeptide |
| U07151 | ARL3 | ADP-ribosylation factor-like 3 |
| AB007144 | DAPK3 | death-associated protein kinase 3 |
| D16431 | HDGF | hepatoma-derived growth factor (high-mobility group protein 1-like) |
| ***Carbohydrate metabolism*** | | |
| D42041 | G2AN | alpha glucosidase II alpha subunit |
| D49737 | SDHC | succinate dehydrogenase complex, subunit C, integral membrane protein, 15kDa |
| ***Others*** | | |
| AB064272 | DPCR1 | diffuse panbronchiolitis critical region 1 |
| AF153605 | AIG-1 | androgen-induced 1 |
| BC008796 | DIPA | hepatitis delta antigen-interacting protein A |
| BC033193 | MGC30208 | hypothetical protein MGC30208 |
| BC000848 | HSA243666 | ubiquitin-conjugating enzyme E2Q (putative) |

## Appendix 9: List of genes which are downregulated after 6hr of WNV infection. (GpIII)

| GB Acc | Symbol | Gene Name |
|---|---|---|
| ***Regulation of cell cycle*** | | |
| AF155235 | NHP2L1 | NHP2 non-histone chromosome protein 2-like 1 (S. cerevisiae) |
| AF480461 | ZAK | sterile alpha motif and leucine zipper containing kinase AZK |
| AL360166 | NME1 | non-metastatic cells 1, protein (NM23A) expressed in |
| D16431 | HDGF | hepatoma-derived growth factor (high-mobility group protein 1-like) |
| ***Others: Regulation of transcription*** | | |
| AB023221 | FBXL11 | F-box and leucine-rich repeat protein 11 |
| AJ224901 | ZNF198 | hypothetical protein BC008631 (with zinc-finger motif) |
| BC005342 | NSBP1 | nucleosomal binding protein 1 |
| BC006143 | MAFG | v-maf musculoaponeurotic fibrosarcoma oncogene homolog G |
| ***Others: G-protein coupled receptor protein signaling pathway*** | | |
| BC008770 | GPR56 | G protein-coupled receptor 56 |
| U07151 | ARL3 | ADP-ribosylation factor-like 3 |
| ***Others*** | | |
| BC033193 | MGC30208 | hypothetical protein MGC30208 |
| BC010436 | MGC16028 | MGC16028 similar to RIKEN cDNA 1700019E19 gene |
| BC014887 | FLJ12270 | hypothetical protein FLJ12270 (with 8 WD domains) |
| M55172 | AGC1 | aggrecan 1 (chondroitin sulfate proteoglycan 1, large aggregating proteoglycan, antigen identified by monoclonal antibody A0122) |
| D50371 | ATP5I | ATP synthase, H+ transporting, mitochondrial F0 complex, subunit e |
| AK023995 | FLJ12442 | hypothetical protein FLJ12442 |
| BC002461 | BNIP2 | BCL2/adenovirus E1B 19kDa interacting protein 2 (anti-apoptotic) |
| AB014548 | SCC-112 | SCC-112 protein |
| AB064272 | DPCR1 | diffuse panbronchiolitis critical region 1 |
| AB018306 | KIAA0763 | KIAA0763 gene product |
| AK001568 | FLJ10706 | hypothetical protein FLJ10706 |
| AJ007714 | AASS | aminoadipate-semialdehyde synthase |
| BC007552 | H3F3B | H3 histone family 3B |

## Appendix 10: List of genes which are downregulated after 18hr of WNV infection. (GpIV)

| GB Acc | Symbol | Gene Name |
|---|---|---|
| ***Nucleic acid metabolism*** | | |
| BC005266 | PRIM1 | primase, polypeptide 1, 49kDa |
| D21851 | LARS2 | leucyl-tRNA synthetase 2, mitochondrial |
| ***Transcription regulation*** | | |
| AB018351 | TOX | thymus high mobility group box protein TOX |
| AF022654 | SHOX2 | short stature homeobox 2 |
| AF507946 | ZNF23 | zinc finger protein 23 (KOX 16) |
| M76541 | YY1 | YY1 transcription factor |
| X89887 | HIRA | HIR histone cell cycle regulation defective homolog A (S. cerevisiae) |
| Y07867 | PIR | Pirin |
| ***Others: Response to external stimulus*** | | |
| BC004168 | FKBPL | FK506 binding protein like |
| M28827 | CD1C | CD1C antigen, c polypeptide |
| M29893 | RALA | v-ral simian leukemia viral oncogene homolog A (ras related) |
| ***Others: transport*** | | |
| AF035621 | KIF3C | kinesin family member 3C |
| AK025790 | RAB6KIFL | kinesin family member 20A |
| AK057954 | NEXN | nexilin (F actin binding protein) |
| AK027543 | MSTP028 | potassium channel tetramerisation domain containing 10 |
| ***Others*** | | |
| M69040 | SFRS1 | splicing factor, arginine/serine-rich 1 (splicing factor 2, alternate splicing factor) |
| AK023971 | FLJ13909 | hypothetical protein FLJ13909 |
| BC005004 | FLJ10156 | hypothetical protein FLJ10156 |
| BC009424 | BANP | BTG3 associated nuclear protein |
| AK023137 | MGAT4B | mannosyl (alpha-1,3-)-glycoprotein beta-1,4-N-acetylglucosaminyltransferase, isoenzyme B |
| AL136910 | FIP1L1 | FIP1 like 1 (S. cerevisiae) |
| BC001376 | SCAMP2 | secretory carrier membrane protein 2 |
| BC001409 | MGC2963 | hypothetical protein MGC2963 |
| AF132940 | LOC51604 | phosphatidylinositol glycan, class T |
| BC003353 | MGC5309 | hypothetical protein MGC5309 |
| AK022617 | IMP-1 | IGF-II mRNA-binding protein 1 |
| AL512725 | MIDN | midnolin |
| AF179786 | ERU162 | Eulemur rubriventer olfactory receptor |
| J03459 | LTA4H | leukotriene A4 hydrolase |
| BC029162 | MGC35440 | hypothetical protein MGC35440 |
| AK056357 | FLJ31795 | hypothetical protein FLJ31795 |
| BC011990 | C20orf30 | chromosome 20 open reading frame 30 |
| BC008440 | LOC51074 | likely ortholog of mouse monocyte macrophage 19 |
| BC003098 | HSD17B4 | hydroxysteroid (17-beta) dehydrogenase 4 |
| BC005004 | FLJ10156 | hypothetical protein FLJ10156 |
| BC003099 | DPH2L1 | DPH2-like 1 (S. cerevisiae) |
| AK025485 | PEO1 | progressive external ophthalmoplegia 1 |
| AK026312 | FLJ22659 | hypothetical protein FLJ22659 |

## Appendix 11: Genes which are upregulated after 18hr of WNV infection. (GpV)

| GB Acc | GenName | |
|---|---|---|
| ***Cellular physiological processes*** | | |
| AB007144 | DAPK3 | death-associated protein kinase 3 |
| AF035444 | PHLDA2 | pleckstrin homology-like domain, family A, member 2 |
| X58531 | LAMA1 | laminin, alpha 1 |
| AB040884 | OSBPL8 | oxysterol binding protein-like 8 |
| AL136635 | RAB1B | RAB1B, member RAS oncogene family |
| M62302 | GDF1 | growth differentiation factor 1 |
| X61118 | LMO2 | LIM domain only 2 (rhombotin-like 1) |
| ***Integral to plasma membrane*** | | |
| AF053004 | WSX1 | interleukin 27 receptor, alpha |
| Y16187 | ECEL1 | endothelin converting enzyme-like 1 |
| AB015355 | SLC11A2 | solute carrier family 11 (proton-coupled divalent metal ion transporters), member 2 |
| BC011403 | GABRA5 | gamma-aminobutyric acid (GABA) A receptor, alpha 5 |
| X13916 | LRP1 | low density lipoprotein-related protein 1 (alpha-2-macroglobulin receptor) |
| ***Others*** | | |
| AB002370 | KIAA0372 | KIAA0372 |
| AB020667 | UBCE7IP5 | likely ortholog of mouse ubiquitin conjugating enzyme 7 interacting protein 5 |
| AB023166 | CIT | citron (rho-interacting, serine/threonine kinase 21) |
| AF249872 | FATE | fetal and adult testis expressed transcript protein |
| AJ006266 | AND-1 | WD repeat and HMG-box DNA binding protein 1 |
| AK023799 | KIAA1271 | KIAA1271 protein |
| AK027372 | FLJ14466 | hypothetical protein FLJ14466 |
| AL136917 | TUFT1 | tuftelin 1 |
| AL390173 | DKFZp547I224 | hypothetical protein DKFZp547I224 |
| BC001430 | RPP20 | POP7 (processing of precursor, S. cerevisiae) homolog |
| BC001784 | HSU15552 | acidic 82 kDa protein mRNA |
| BC032512 | MGC15631 | hypothetical protein MGC15631 |
| M26167 | PF4V1 | platelet factor 4 variant 1 |
| --- | --- | Protein containing four C2H2 type zinc finger domains |
| --- | --- | Protein containing 9 TPR (tetratrico peptide repeat) domains |

## Appendix 12: Genes which show upregulation only at 6hr during WNV infection. (GpVI)

| GB Acc | Symbok | Gene Name |
|---|---|---|
| ***Endoplasmic reticulum*** | | |
| AF132048 | RTN4 | reticulon 4 |
| BC001549 | EBP | emopamil binding protein (sterol isomerase) |
| U00968 | SREBF1 | sterol regulatory element binding transcription factor 1 |
| BC028679 | KIFAP3 | kinesin-associated protein 3 |
| AF215935 | P44S10 | proteasome regulatory particle subunit p44S10 |
| AK001135 | SEC23IP | SEC23 interacting protein |
| ***Cytoskeletal functions*** | | |
| AK023385 | TPM3 | tropomyosin 3 |
| AK023385 | TPM4 | tropomyosin 4 |
| U87166 | ABI1 | abl-interactor 1 |
| BC028679 | KIFAP3 | kinesin-associated protein 3 |
| ***Mitochondrion*** | | |
| AB035863 | SUCLA2 | succinate-CoA ligase, ADP-forming, beta subunit |
| AF020038 | IDH1 | isocitrate dehydrogenase 1 (NADP+), soluble |
| BC000187 | COX6C | cytochrome c oxidase subunit VIc |
| BC016350 | ATP5F1 | ATP synthase, H+ transporting, mitochondrial F0 complex, subunit b, isoform 1 |
| M75883 | SCP2 | sterol carrier protein 2 |
| Z35093 | SURF1 | surfeit 1 |
| U31930 | DUT | dUTP pyrophosphatase |
| ***Others: Intracellular transport*** | | |
| BC015179 | SNX3 | sorting nexin 3 (involved in protein trafficking) |
| BC001255 | NCBP2 | nuclear cap binding protein subunit 2, 20kDa |
| BC019944 | SLC20A1 | solute carrier family 20 (phosphate transporter), member 1 |
| AK001135 | SEC23IP | SEC23 interacting protein |
| X57347 | YWHAQ | tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein (14-3-3), theta polypeptide |
| ***Others: Transcription regulation*** | | |
| BC013998 | PITX2 | paired-like homeodomain transcription factor 2 |
| L03411 | RDBP | RD RNA binding protein |
| X70394 | ZNF146 | zinc finger protein 146 |
| ***Others: Protein metabolism*** | | |
| AF283772 | RPL15 | ribosomal protein L15 |
| AK000734 | FLJ20727 | ubiquitin specific protease 47 |
| BC000848 | UBE2Q | ubiquitin-conjugating enzyme E2Q (putative) |
| BC009839 | SKP1A | S-phase kinase-associated protein 1A (p19A) |
| M64174 | JAK1 | Janus kinase 1 (a protein tyrosine kinase) |
| U73824 | EIF4G2 | eukaryotic translation initiation factor 4 gamma, 2 |
| ***Others: Signal transduction*** | | |
| BC004186 | GNB1 | guanine nucleotide binding protein (G protein), beta polypeptide 1 (for signal transduction) |
| AF061243 | MP1 | pitrilysin metalloproteinase 1 |
| BC005219 | NRAS | neuroblastoma RAS viral (v-ras) oncogene homolog |
| ***Others*** | | |
| AB002370 | KIAA0372 | KIAA0372 (contains multiple tetratrico peptide repeat domains) |
| AB011173 | KIAA0601 | KIAA0601 protein (flavin containing amine oxidase family) |
| AF226052 | FLJ20432 | hypothetical protein FLJ20432 |
| AF272973 | EAF1 | ELL associated factor 1 |

| | | |
|---|---|---|
| AF342736 | BBS2 | Bardet-Biedl syndrome 2 |
| AK024552 | PAIP2 | hypothetical protein BC007764 |
| AY007103 | HSPC195 | CXXC finger 5 |
| BC003409 | OCIA | ovarian carcinoma immunoreactive antigen |
| BC005004 | FLJ10156 | hypothetical protein FLJ10156 |
| BC005968 | APACD | ATP binding protein associated with cell differentiation |
| BC008070 | HSPC182 | HSPC182 protein |
| BC008440 | LOC51074 | likely ortholog of mouse monocyte macrophage 19 |
| BC008796 | DIPA | hepatitis delta antigen-interacting protein A |
| BC011990 | C20orf30 | chromosome 20 open reading frame 30 |
| BC017701 | DKFZp434L142 | hypothetical protein DKFZp434L142 |
| BC022508 | GPM6A | glycoprotein M6A |
| X76534 | GPNMB | glycoprotein (transmembrane) nmb |
| BC025296 | ODC1 | ornithine decarboxylase 1 |
| U81607 | AKAP12 | A kinase (PRKA) anchor protein (gravin) 12 |